

DOCUMENT RESUME

ED 151 736

CS 003 963

AUTHOR Wiener, Morton; Shilkret, Robert
TITLE Complexity in Auditory and Graphic Language. Final Report.
INSTITUTION Clark Univ., Worcester, Mass.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
BUREAU NO 4-470
PUB DATE 15 Dec 77
NOTE 208p.
EDRS PRICE MF-\$0.83 HC-\$11.37 Plus Postage.
DESCRIPTORS *Child Language; *Complexity Level; *Comprehension; Context Clues; Language Research; Language Usage; Measurement Techniques; Reading Comprehension; *Reading Research; Redundancy; *Speech Communication; Syntax; *Written Language

ABSTRACT

Starting with a model for explaining comprehension and noncomprehension of verbal material in terms of a match/mismatch principle, this project developed a scale of language usage and explored hypotheses about how comprehension may become possible if a child does not now comprehend some particular oral or written text. Eight separate reports are included. They discuss the development and testing of the scale; the likelihood that a child will comprehend a text of lesser, the same, or greater complexity than that usually used by the child; the possibility of increasing the comprehension of complex materials by introducing redundant features; the effects of "associativity" as a form of redundancy and as a factor in comprehension; the language features of experienced and inexperienced caretakers of nursery school children; and the effects of these features upon the children's language production. (Author/AA)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

FINAL REPORT

Complexity in Auditory and Graphic Language

National Institute of Education Project No. 4-470

Morton Wiener, Ph.D., Principal Investigator
Robert Shilkret, Ph.D., Co-principal Investigator

Clark University
Worcester, Mass. 01610

December 15, 1977

The research reported herein was performed pursuant to a grant with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

Summary

Starting with a model for explaining comprehension and non-comprehension of verbal material in terms of a match-mismatch principle, an attempt was made to develop a scale of language usage to assess a match, and to explore hypotheses about how comprehension may become possible if a child does not now comprehend some particular oral or written text.

A scale of complexity of language usages is detailed and tested. The likelihood that a child will or will not comprehend a text when it is written with lesser, with the same, and with greater complexity than that used by the child was tested. The possibility of increasing the comprehension of complex materials by introducing redundant features in an utterance was explored. In addition, effects of "associativity" as a form of redundancy and as a factor in comprehension was examined. Last, the language features of experienced and inexperienced caretakers of children -- assessed for degree of complexity and redundancy -- was explored when giving information to nursery school children and the effects upon the language production of the children examined.

ACKNOWLEDGMENTS

We are grateful for the invaluable help of our co-authors, colleagues and students who made most, if not all, of this work possible:

The study "A scale of complexity of language: An application to the utterances of 4th, 6th & 8th grade children from middle and lower socioeconomic neighborhood schools" is based in part on a Master's Thesis submitted by Maureen Rubano to the Department of Psychology, Clark University, 1976. We thank Lynda Godkin, Suzannah Heard, Gail Hornstein, Laurie Lockwood, Craig Wiener and Gary Yablick who served as scorers and/or assisted with other aspects of this research.

The study "Language complexity of three age groups of middle class children for two kinds of stimuli with similar content" was completed as an Independent Work Project at Mt. Holyoke College, 1977, by Nancy Haley, Cathy Hill and Mimi Hayden. We thank Maureen Callanan, Anne Claunch and Lynda Godkin who assisted us in the scoring and analysis of the data of this study.

The report "Reading difficulty of three levels of text complexity for readers who differ in the complexity of their verbal production" is based in part on a doctoral thesis being written by Rick Schnable in partial fulfillment of his Ph.D. at Clark University.

We thank Anne Claunch and Maureen Callanan who also assisted in this scoring.

The report of the "Effects of redundancy and intonational stress in children's comprehension of temporally ordered events" is based in part on an Honors Thesis submitted by Marion Viets, 1975, Mt. Holyoke College.

The report, "Comprehension of sentences expressing three events and the temporal relations between them" is based in part on an Honors Thesis, by Patricia Arns, submitted to Mt. Holyoke College, 1976.

The study "Knowledge of referent and comprehension of language" is based in part on a Master's Thesis by Judith Cormier, submitted to Clark University, 1976. We thank Gail Hornstein, Dennis Wixon and Ilene Spiro for their invaluable aid in scoring the verbal protocols.

The report of the study "Modification in three groups of adults' speech to children and effects on children's language" is based in part on an Honors Thesis by Lynnwood Andrews, Mt. Holyoke College, 1976. We thank Lynda Godkin who assisted in the collection of data and Maureen Callanan, Anne Claunch and Suzannah Heard who assisted in the scoring and analysis of these data.

We gratefully acknowledge the contribution of our other colleagues

and students who participated in our ongoing project meetings - Shannon Devoe, Jo Ann Bennett, Andrea Hart, Gail Hornstein, Dennis Wixon and Gary Yablick -- and helped us to clarify our concepts, our scoring criteria and our methodological procedures. In addition, Dennis Wixon was invaluable in helping us with statistical problems.

This acknowledgment would be less than complete without an expression of thanks and appreciation to the School systems, staff, and principals who made it possible for us to complete all of the work reported as part of this project:

(1) Shrewsbury, Mass. School system, Irving Agard, Superintendent; Robert Hutchinson, Curriculum Coordinator, Sooren Kachadoorian, Principal, Beale School and John Bohdiewicz, Principal, Ward School.

(2) Frances Kerr, Director of the Gorse Child Study Center of Mt. Holyoke, Mass.

(3) The Ayer, Mass. School System, Neville Markham, Superintendent, Dorothy Kron, Reading Supervisor, and Loretta Eliades, Principal, Page Elementary School.

(4) Worcester, Mass., School System, John Connors, Superintendent; Esther Smith, Principal, Flagg Street School; James Duffy, Principal, Forest Grove Jr. High School; John Monfredo, Assistant Principal, Woodland Street Annex School; and Bruce Wells, Principal, Worcester East Middle School.

(5) South Hadley, Mass., School System, Gerald Hoar, Superintendent; Joan Tapir, Principal, Center School; Gerald Greaney, Principal, Plains School; and Edgar Noel, Principal, Intermediate School.

To the many, many teachers in all of these schools who cooperated so very well and tolerated our intrusions into their classroom, we express our thanks.

Finally, none of these studies would have been possible without the participation and cooperation by the children and their parents. We owe them a debt of gratitude.

This project derives from a general framework which explores successes and failures of comprehension of both auditory and graphic language instances in terms of the typical patterns of communication used by a communicator and the communication requirements as an instance to be comprehended by that communicator. That is, the basic model proposes that comprehension requires a match between the patterns of communication an individual or group brings to a communication instance -- however measured -- and the pattern of communication in the instance to be comprehended assessed independently, but by the same scale. The assumption is that instances of success or failure of comprehension can be understood in terms of a comparison of the "features" of language used and/or responded to by a communication and the requirements of the "features" present in the language instance to be comprehended. A brief discussion of the model and some of its assumptions may be necessary before reporting the particular investigations undertaken as part of this project.

Within a language community, the behaviors which are taken as "comprehension" seem to be a commonplace occurrence for most individuals and for most communication instances. One rarely considers comprehension as a problem except for children or for others who are usually thought of as defective or deficient. Indeed, if language comprehension occurred for all adults in all instances, one might question whether comprehension would arise as a general problem for psychology rather than as a problem for pedagogy or for individual differences in the sense of differences in "intelligence." That is, the general problem "Under what conditions is comprehension possible?" becomes apparent only in those instances when an adult assumed to be efficient in language usage "miscomprehends" or fails to comprehend a given communication (i.e., and the failure cannot be attributed to the comprehender as if he were "defective" in some way). However, if one considers this issue carefully, it is clear that almost all members of a language group have difficulty in some kinds of language instances (e.g., a psychologist reading a physics book).

In this context, it is evident that any consideration of language behavior will indicate that a given instance can be characterized by some statement of its "complexity" (e.g., semantic components, syntactic form; sentence length) and the "difficulty" for a given individual (or group) relative to other language instances. It would seem possible, in principle at least, to evolve a scale to compare and contrast all instances of communication in terms of complexity and predictive difficulty for any specific individual. Were such scaling possible, it would follow that for any individual what is taken as comprehension behavior would be seen to occur only for some set of instances; that is, for those instances which represent only some portion of the total possible range of any such scale.

Despite the possibility of comparing any language instance with other language instances and thereby deriving a continuum of "complexity" and apparent predictive difficulty, language users typically seem to respond in an "either-or" manner. If, as we hold, a scalar notion of language usages is meaningful over the whole range of instances despite the "either-or" comprehension of language users, then it would seem necessary to posit some characteristics of a language event (e.g., in speech, the intonational contour)¹ which, when present, facilitate the

¹In the discussion in this section, we will use the intonational contour of speech as an example of a set of features which, we propose, may "compensate for" increases in complexity; later, we will discuss this same set of features, but instead characterize them as "redundant with" other aspects of the language. Whether a feature is "compensatory" or redundant is dependent upon the other language features present -- for particular subjects who may or may not respond to the available features (as determined by independent judgments). That is, in the most general sense, as we use the terms, a set of features which at one time may be said to "compensate for" an increase in complexity may at some later time come to be labelled as "redundant with" the more complex forms. Thus, for us, the assessment of "redundancy" or "compensation" implies some independent measure of the individual's (group's) performance without a particular set of features. It is only when there is no performance decrement (or increment) without this set of features that they can be said to be "redundant with" or "compensating for" other aspects of the language instance. One implication of this kind of distinction is the requirement that the designating of a particular set of features as redundant or compensatory can only be assigned in an independent analysis for different individuals, groups, or contexts. For example, complexity, which in principle is scalable and should be reflected in terms of performance difficulty may for particular testing instances not show this effect. Lack of experience with all of the syntactic forms, as is the case with beginning speakers may "dissolve" the scale since all forms might be of such great difficulty that any differences between them are so small as to be indistinguishable. Concomitantly, for individuals with a great deal of familiarity with all the forms of the scale (experienced language users, for example) differences between items of the scale may not be manifest, since all forms may be extremely "easy" as a function of previous mastery. It would not be expected that "removing" intonational features would make the input significantly more difficult than it already is for inexperienced speakers, nor would highly experienced speakers find the loss of this additional information, for them redundant, any hindrance. Early during acquisition of language, the added information of intonational features may be critical

(Footnote 1 continued)

possibility of comprehension, and in this sense "compensate" for an increase in the complexity of the forms of the language. Given (1) a conceptualization that language forms can be ordered, and (2) an analysis which makes it possible to denote features of the stimulus which may "compensate" for the increased complexity, it should then be possible to analyze and predict instances of comprehension and non-comprehension -- both for cases when compensatory features are "absent" as well as for cases when these features are "present," as long as we also have some independent indicator of the range of complexity a given individual (or group) has mastered.

There are at least two kinds of instances of non-comprehension which can be considered: within-individual (or group) comparisons and between-individual (or group) comparisons. First, an individual (or group) may fail to comprehend certain instances within a language, given that the same individual (or group) does comprehend other instances within that language. Second, an individual (or group) may fail to comprehend certain instances of language, given that another individual (or group) succeeds in comprehending those same instances in this language. While each of these two kinds of non-comprehension may be considered independently, our focus of concern in this project is for those individuals and for those instances where these two kinds of non-comprehension appear to co-occur. That is, we do not at this moment consider it a "problem" to be accounted for if a given individual who has experienced only English fails to comprehend French.

Much of the work within the broad field of language comprehension can be seen as derived from one or the other of two quite different assumption models. (1) On the one hand, much recent work has focused on what might be called "characteristics of language" exemplified by the "psycholinguistic" approach. The general procedure of this approach seems to be to

¹for comprehension and for the learning of new grammatical forms. With experience with particular forms, and mastery of them, intonational features may, in fact, become "redundant" rather than "compensatory," and, therefore, not as critical for comprehension. From the above argument, it should be possible to construct scales of complexity, which in terms of a performance measure are "too easy" for one population, "too difficult" for another population, but yet "scale out" for a third intermediate group. It is only when such scaling is possible, as in the intermediate group, that we would expect to be able to test for the posited compensatory effects of intonational features. In sum, compensation or redundancy are terms which can be considered only in some specifiable context and the basis for designating an item as compensatory or redundant only in relationship to other language features, for particular language users.

select a group of subjects, assumed to be competent and equivalent language users, who are assessed in experimental tasks requiring them to perform with a variety of language instances, ordered by some concept of complexity. Since the subjects are assumed to be identical, performance variability is assumed to derive from differences in the stimuli alone (i.e., between language instances). (2) On the other hand, other work within this field has focused on differences between language users, exemplified by much of the work in the area of reading difficulties, and more generally the recent "sociolinguistic" approach. Often the assumption within this kind of approach appears to be that variability of language performance is to be attributed to individual differences alone (e.g., the characterization of an individual as a "poor reader" with no specification of the material which is or is not comprehended by that individual). Certainly each of these two approaches is valuable; however, adopting either one or the other of them seems to lead to logical or empirical difficulties which are often not apparent until the assumptions are made explicit. For example, recent work which has attempted to adopt a "psycholinguistic" approach to "reading problems" often also includes a concern with assessing a range of variability both within and between individuals. More often in these kinds of approaches the concern appears to be only with one particular group (i.e., those labelled "poor readers"), without apparent recognition by the investigator that the same kinds of concerns can be applied to within individual instances of comprehension or non-comprehension.

As a general approach, our past investigations and those proposed here focus on exploring and specifying the relationships between what can be called (1) response-availability characteristics of comprehenders, conceptualized as "patterns of behavior" an individual brings to the task, and (2) characteristics of the stimulus to be comprehended. By adopting such an "interactionist" view,² we hope to avoid some of the

²We do not regard our approach to be an "eclectic" one; eclecticism implies principles of selecting between each of two (or more) different assumption models, in the hope of achieving the "best of both possible worlds." Since an "eclectic" approach rarely makes explicit the principles of selection, it is most often less principled than post hoc. Eclecticism in general usually refers to a combination of two or more viewpoints; more particularly, the term usually refers to the combination of two or more implications (e.g., plans of action) deriving from different viewpoints. Often, the different points of view themselves are contradictory, and thus, logically incompatible. In contrast, when the assumptions underlying different viewpoints are made explicit, are found to be incompatible and a different model is then proposed, "reinte-

(Footnote 2 continued)

conceptual difficulties of (1) dealing with comprehension as the "decoding" of material in the stimulus itself, as if independent of a particular comprehender (the metaphor of "empiricism;" "naive realism"), and (2) dealing with comprehension as if it were "in" the comprehender, that is, as if independent of the context to be comprehended (the metaphor of "nativism;" "innatism").

While the focus of our point of view is on the relationship between the general class "event-to-be-comprehended" and the general class "comprehender's usage," we assume that each may be assessed independently, at least from the point of view of an ideal observer. We also assume the possibility for positing an "ideal" scale or ordering the complexity of a language instance in terms of logical relations in the context of the language form. The construction of such a scale assumes that event A' is taken to be more "complex" than event A, if A' includes the discrimination of A, as well as at least one additional discrimination, combination or relation. For example, if an event is a correlation or coordination of two components (X and Y), this instance is taken to be less "complex" by such a scale than, for example, the conditionality of the same two components (X if Y); conditionally assumes correlation as well as an additional relation, i.e., the necessary condition for X. In the same way, other logical relations could be "scaled" on such a measure of complexity.

Given this kind of ideal, logical scale, different verbal constructions of events could be assessed in terms of their logical complexity. Concomitantly, using the same kind of scale, individuals could be ordered in terms of the degree of logical complexity of their construction of events in different forms of behaviors (e.g., production vs. comprehension; differences between different domains of performance, such as language vs. action). Given such independent assessments of an individ-

²gration" would seem to be a more appropriate characterization than "combination" implicit in "eclecticism." It is our assumption that any statement of "comprehension" must include both a statement of the comprehender as well as a statement of the material-to-be comprehended. At root, this kind of interactionist assumption constitutes a different "world hypothesis" (Pepper, 1942) and is not derivable from either of its components alone or in combination. This approach suggested here is analogous to some earlier work in perception (Wiener and Schiller, 1960; Kempler and Wiener, 1963; Wiener and Kleepsies, 1968) which attempted to show that perceptual activity (e.g., identification) can best be accounted for in terms of what a perceiver brings to a particular set of stimuli, requiring a particular kind of response, rather than noting characteristics of the perceiver alone, or the stimulus alone.

ual's particular forms of construing events by this scale as well as an ordering of language utterances by this scale, one could predict comprehension of, or "adequate" performance in a given instance by a given individual, as well as predict differences in comprehension or performance based on the kind of "mismatch" of forms available for an individual for particular kinds of instances.

It should be noted that this kind of principle of ordering (i.e., incorporation of elements, grouping of elements, and relations) is conceptualized as an "ideal" scale (Kaplan, 1966; Toulmin, 1953), and is used solely as a heuristic device in an attempt to order coherently and economically, and predict phenomena which it describes. Thus, such a logical scale is not considered as "inherent in" the events themselves, the comprehender's manner of constructing events, or even the interaction between the two. It may well be that under certain conditions, for certain comprehenders and/or for certain events, the empirically observed order of "difficulty" (i.e., as indicated by a performance measure, such as those to be discussed below, some of which we and others have used in previous investigations, others of which we hope to develop as part of the program of research proposed here) may be found not to correspond with the posited "ideal" order of complexity which we expect. With such empirical findings, we would be required to re-explore our original scale and consider modification and reconsideration of the "logic."

A second kind of issue in the use of the proposed "ideal" ordering derives from the assumption of ordinality ingredient in the way we have defined the scale. While we posit an increase in complexity with the inclusion of one further discrimination requirement, and each addition is assumed to increase complexity, additivity (interval scale) is not assumed in our derivation of an ordinal scale. That is, while the same differences in scale "values" may be obtained at different points in the scale, we do not hold that the degree of change need be the same at the two points -- they are only different in an ordinal sense (i.e., greater or lesser than).

While we have outlined the idea of a scale as if it were possible to derive such a scale completely independently of any empirical evidence, and as if we were concerned only with testing the empirical tenability of such a scale once derived, we, in fact, see the idea of positing an "ideal scale" as an heuristic device. First, the articulation of particular aspects of language behavior which may be of importance in a scale of complexity will be accomplished in the context of testing specific hypotheses derived from our previous work, and the work of others. Second, as mentioned above, instances of non-correspondence may be found between an expected order of difficulty and an observed order of difficulty in the linguistic performance of a particular group under particular circumstances.

Within our general methodology for the use of this kind of heuristic scale, when such non-corresponding instances occur, requires that we look for other co-occurring variables which could account for the discrepancy of the ideal predictions from the empirical observations.

In contrast to many other approaches using complexity (e.g., using Chomsky's transformational grammar in deriving an ordering of difficulty) to explore language comprehension behavior, the model proposed here requires that any complexity measure include at the very least: (1) syntax (i.e., principles for combining components and associated variations in sentence form); (2) associativity among components similar to Rosenberg's (e.g., 1968) concept of associativity and Perfetti's (e.g., 1969b) "lexical density," but not restricted to the sentence as the unit of analysis but rather the larger communication matrix; (3) number of discrete components, or length; (4) degree of correspondence between sequencing of components in "real time" and the temporal-spatial ordering of the represented event; (5) implies requirements of relational terms (e.g., temporal-spatial sequence vs. causality; coordination vs. conditionality; (6) explicitness of the communication features available to indicate variations in all of the above.

We begin with the further general assumption that (with the exception of perceptual impairments for a particular individual) the ease or difficulty of language comprehension is, in principle, equivalent in the auditory and visual modalities. Ostensible "superiority" of comprehension of auditory forms (compared with comprehension of visual forms), we hold, is usually attributable to lesser complexity of the material to be comprehended, less demanding criteria of "comprehension" or availability of some specific auditory cues not available in the visual form. Thus, if stimulus complexity and comprehension criteria are equivalent in the auditory and visual modes and the auditory form is still easier to comprehend than the corresponding visual form, we would expect that there is available in the auditory form an explicit cue for the organization of the material which is not available in the ostensibly equivalent graphic form. That is, spoken English includes an international contour consisting of "melodic features" (e.g., patterns of stress, pitch, hesitation) which are neither fully explicit nor articulate in the graphic form of English.

This kind of analysis would imply that the degree to which a particular language has explicit graphic forms which correspond to vocal organizational features of the language, the easier such a language will be to learn to read (to the extent that writing can be viewed as a notational system for representing the spoken form of language). Besides the degree of explicitness of a notational system for designatory organization (as to some melodic contours), different written languages may also vary in terms of (1) phonemic explicitness, that is,

the extent to which there is a different graphic form for each phonemic variation, and (2) phonemic articulateness, that is, the extent to which each grapheme has the same phonemic realization in a variety of contexts. Written English is neither fully explicit nor fully articulate (e.g., a given grapheme may have several phonemic realizations depending on the context of occurrence -- as, for example, in the case where the initial vowel sound is dependent on the presence or absence of a subsequent "silent e" -- "cut" versus "cute"). Of relevance here is Makita's (1968) exemplification of the high degree of correspondence in Japanese (as compared with Western languages) between graphic forms and phonemes, and the concomitant lesser difficulty of reading for Japanese children, compared to Western children at equivalent points during acquisition. While Japanese seems both more explicit and more articulate than English, notational systems with even more overall explicitness and articulateness might be possible (e.g., the notational system of formal logic; musical notation).

Given a less than fully explicit notational system in writing, the framework of the present research implies that part of learning to read English must involve learning to organize the input appropriately on the basis of some cues which are available; in this sense, learning to read English involves learning to impose organization on a visual input which is neither fully explicit nor articulate, an organization which does occur in auditory forms. In this context, we hold that the accomplished reader of English has mastered a complex set of such "rules" for organizing the visual input. Further, for such an accomplished reader, it would not be expected that making the visual input more explicit and/or articulate (e.g., by providing additional visual cues) would facilitate comprehension. However, it would be expected that for some individuals, labelled "poor readers," the difficulty is in organizing the visual material appropriately. In this sense, the difficulty is broader than one of reading alone, and it would be expected that these individuals would show evidence of equivalent difficulty in auditory language performance when organization "cues" in speech (i.e., certain intonational features) are experimentally manipulated.

It is important to note that our approach makes no assumption concerning which particular set of attributes is "developmentally prior," and thus serves as the set of "redundant" features in making change possible. In this particular instance, we regard this question as an empirical one. However, there is considerable evidence at least for some constructions -- the syntactic form of the question, for example -- that the intonational contour is developmentally prior to the syntactic form. Werner Leopold seems to have been the first specifically to point out the importance of intonational features in speech development (see Werner and Kaplan, 1963; Ingram, 1971), although his observations have been confirmed by more recent

investigators (e.g., Weir, 1962). More recently, evidence has appeared to suggest that infants (8 mos.) discriminate certain intonational contours (Kaplan and Kaplan, 1971; Morse, 1971); that children imitate (e.g., Miller and Ervin, 1964) and produce spontaneously (e.g., Gruber, 1967; Kagan, 1964; Gleitman and Shipley, 1971; Weeks, 1971) the intonational features of utterances prior to the syntactic features with which they co-occur. From this kind of evidence, we would argue that the syntactic specificity, at least for some constructions, is the "redundant" information; that is, a particular syntactic form may be "redundant" with an earlier intonational contour ("old" form).

While most of our previous work has focused on melodic features, as a source of redundancy, other kinds of examples could be cited, some of which will be explored in the research reported here. For example, we would argue in general that the more explicit and articulate language forms which are available to express particular relationships, the easier such relationships will be to "comprehend" and the earlier they will be "acquired." Evidence in support of the first part of this kind of hypothesis (the "efficiency" expectation) was provided recently by Markman (1973). She found that it was much easier than previously thought for young children (1st graders) to answer correctly the classical Piagetian class-inclusion problems when the relationships were indicated by non-ambiguous (i.e., explicit and articulate) language forms. Specifically, it was found that her subjects would make part-whole comparisons which would indicate the "presence" of class-inclusion concepts with pictures of "families" as stimuli. Families were used as stimuli in that the subparts of the family (parents and children) are not labelled "family." Thus, in this case, where there were mutually-exclusive terms referring to a superordinate class and its subclasses, "class-inclusion" was possible for children who had already mastered those labels.

Up to this point, we have outlined a model which assesses current performance of subjects for particular tasks. We hope, given a model which can assess the match-mismatch characteristics, now to explicate a strategy for inducing change so that subject will respond to a more complex form of language than the one to which s(he) now responds.

In the most general sense we regard two criteria as essential for the possibility of "change" from an "old" form to a "new" one. (1) The two forms (i.e., "old" and "new") must be ordered in the way we have discussed the use of an "ideal" scale -- that is, it must be possible to assess independently both forms, and the forms must be related ordinally (i.e., any "new" form includes "old" form, plus some additional components or relation). (2) The transition from "old" to "new" forms, the "change" itself, can be accomplished if there are co-occurring characteristics of the stimulus which are redundant

with both forms. Once the "new" form has been mastered with inclusion of the "old" redundant form(s), change to the more complex form without the old forms can become possible. In the most general sense, then, the following can be used as a schematic representation of the conditions for "change."

1. "Old" form
2. "Old" form together with "new" form³ (redundancy)
3. "New" form without "old" form

It should be noted that we regard the two broad criteria outlined above as essential for "change." Whether these criteria are also sufficient for "change" is, from this point of view, an empirical question. That is, given that the two criteria are met, we would argue that change is possible, whether or not change occurs must be assessed in terms of particular individuals (or groups) in particular contexts. However, as a general methodology, we hypothesize that these criteria are sufficient for "change" in any particular study. As before, when no evidence of "change" is found, our model would suggest a careful re-analysis of those characteristics (e.g., "redundant" features) which may be introduced to facilitate "change."

It should also be noted that this kind of model of "change" hopefully avoids the problems of many developmental models which presupposes "conflict" as a prerequisite for change (e.g., Turiel, 1969). The concept of "mismatch" as conceived here does not imply the necessity that the comprehender indicate that he is "aware" of the "mismatch." In fact, we argue that it is reasonable to assume that the comprehender always "comprehends the world in his own terms." This assertion does not deny that in some cases, (s)he may well indicate his/her "awareness" of a "mismatch" between his/her "comprehension" and the particular requirements of the context at that time. However, we would assume that this awareness occurs only subse-

³This procedure implies that what is to be taken as "new" for a given individual (or group) may be determined only in conjunction with some independent assessment of his prior learning (i.e., the set of "old" forms he has mastered). In one sense, this criterion is similar to the one of "inner logic" between "developmental stages" Loevinger (1969) has discussed regarding "Piaget's developmental theory. However, by adopting a somewhat "weaker" criterion for assessment of "change" (i.e., any additional component or relation) we hope to avoid the possible logical and empirical difficulties which may ensue from assuming a particular order is universal, as if "in" the organism (see earlier discussion of this issue). Additionally, we hold that this kind of model has allowed for, and will continue to allow for, more specific predictions concerning individual differences.

quent to some feedback upon a prior "comprehension." In this special kind of instance, the feedback information would be expected to include some component(s) which is redundant with the individual's original "comprehension" of the situation (the "old") as well as some "new" information. From this point, our general analysis of change would apply even in these special instances. More generally, the "mismatch," as conceived here, is solely a statement by an observer, and derived from independent analysis of the individual and the task in terms of the same "ideal scale."

This project then, is an attempt to (1) explicate a scale of complexity of language usages which is used to assess the syntactical and logical complexity of an utterance; (2) assess the likelihood of a child comprehending a text when it is written with lesser, with the same, and with greater complexity than that used in the typical communication patterns by that child; (3) explore the possibility of increasing the comprehension of complex materials by introducing a redundant feature in an utterance, a feature which is less complex than the semantic components to be comprehended -- in this case tonal stress patterns; (4) test the proposition that more complex material may be comprehended after the child has an occasion to "learn" the more complex form in the context with a redundant less complex semantic form -- in the instance explored, a semantic feature (that is, a child will learn to respond to "before" as a temporal maker, if the child first has an opportunity to respond systematically to "before" in a context where the temporal ordering is also marked by the term "first," e.g., "Do X first, before you do Y," followed by a test of "Do L before you do M"); (5) explore the effects of "associativity" as a factor in comprehension; and (6) explore the language features of "experienced" and "inexperienced" caretakers of children when they tell a story to pre-elementary grade school children (ages 4 years 0 months to 5 years 9 months).

We are reporting the following studies which explore these concerns:

1. "A Scale of Complexity of Language: An Application to the Utterances of 4th, 6th and 8th Grade Children from Middle and Lower Socioeconomic Neighborhood Schools." Maureen Rubano, Morton Wiener and Robert Shilkret. Portions of this paper were presented at the American Psychological Association meeting, San Francisco, August 1977. This study reports a set of principles and exemplifies a scale for the assessment of verbal complexity. The scale is then used to (a) assess the degree of complexity on the verbal productions of children from middle and lower socioeconomic class backgrounds; (b) to compare the verbal production of these children for three different tasks posited to differ in complexity; and (c) compare the verbal productions for children in three grades -- 4th, 6th and 8th.

2. "Language Complexity of Three Age Groups of Middle Class Children for Two Kinds of Stimuli with Similar Content." This study compares the degree of complexity evident in the verbal productions of children of three age groups from middle class backgrounds to stimuli which differ in the form of pictorial

representation, that is to a movie versus 13 discrete frames of the same "story."

3. "Reading Comprehension Difficulty of Three Levels of Text Complexity for Readers Who Differ in the Complexity of their Verbal Production." Rick Schnable, Morton Wiener and Robert Shilkret. This study explores the degree to which a child comprehends a text when it is written with lesser, the same or greater complexity (as assessed by the Rubano, Wiener & Shilkret scale) than the level of verbal complexity evident in the child's own verbal productions.

4. "Effects of Redundancy and Intonational Stress in Facilitating Childrens' Comprehension of Temporally-ordered Events." Robert Shilkret, Marian Viets and Morton Wiener. Portions of this paper were presented at the Annual Meeting of the Linguistic Society of America, San Francisco, 1975. This study explores the use of redundant forms (first, second, last) and intonational stress, as a second type of redundancy, in facilitating preschooler's comprehension of temporally ordered events.

5. "Effects of Redundancy in Facilitating a Child's Comprehension of Verbal Forms of Differing Complexity: Tested in a "Before" and "After" Paradigm." Morton Wiener, Judith Cormier and Robert Shilkret. This study explores more systematically children's comprehension of a series of complex verbal forms and the facilitating effects of redundancy on comprehension when (a) redundancy occurs verbally in transition sentences presented systematically to the child and (b) when redundancy is represented by "associativity" in the content of the tasks.

6. "Comprehension of Sentences Expressing Three Events and the Temporal Relations Between Them." Patricia Arns and Robert Shilkret. Paper presented at the 48th Annual Meeting of the Eastern Psychological Association, Boston, April, 1977. This study -- using adult subjects -- explores the "difficulty" of comprehending sentences which differ both in their complexity as well as the degree of subordination used in the utterance.

7. "Knowledge of Referent and Comprehension of Language." Judith Cormier, Morton Wiener and Robert Shilkret. Portions of this paper presented at the 48th Annual Meeting of the Eastern Psychological Association, Boston, April, 1977. In this study we have explored associativity and/or prior knowledge of events as a source of difficulty for comprehension of a text.

8. "Modifications in Three Groups of Adult's Speech to Children and Effects on Children's Language." Robert Shilkret, Lynnwood Andrews and Morton Wiener. Experienced and inexperienced child caretaker's speech to children were analyzed for (a) degree of complexity of their utterances, (b) for the presence and/or absence of redundant intonational patterns and (c) the relationship of differing degrees of complexity by an adult to the subsequent production of verbal materials by the child.

While all of this work reported here is concerned with "basic" issues of learning and some factors which may facilitate learning, we believe that were the posited relationships found to be tenable in these and further experimental work, such findings would have marked implication for the classroom situation itself. We have tried to specify some of these implications in the several separate reports which make up the summary of the work of this project.

A Scale of Complexity of Language: An Application
to the utterances of 4th, 6th and 8th grade children
from middle and lower socioeconomic neighborhood schools¹

Maureen Rubano and Morton Wiener and Robert Shilkret
Clark University Mount Holyoke College

This study was an attempt to articulate and test a proposed scale for ordering terms in English which relate events to each other. The proposed scale was used to explore: (1) age-related differences in the use of these relational terms; (2) the use of such terms in descriptions of differing visual stimuli; (3) differences in usages of these terms by children from different socialization backgrounds.

Wiener and Shilkret (1974) have begun to articulate a set of principles for analyzing and comparing language instances in terms of a posited "ideal" scale of "complexity." Their model views comprehension (or non-comprehension) of communication instances as a function of a match (or mismatch) "between what can be called (1) response availability characteristics of comprehenders...and (2) characteristics of the stimulus to be comprehended." The scale of complexity used in this study is to be regarded only as one possible subset of a more exhaustive scale which they propose.² The general model describes complexity as a function of the number of discriminations inferred from an utterance. In this scale, the relevant discriminations are of simultaneous relations between predications explicitly indexed

this definition warrants further elaboration, as each bears on the principles of the proposed scale.

The unit of analysis (the boundaries within which the number of discriminations are summed) is a "thought unit." A thought unit is defined as a subject and predicate (as defined in conventional grammar), or a number of subjects and predicates conjoined by appropriate connectives other than and.³ Within a single thought unit there may be only one subject-predicate pair, or there may be several pairs with a connective conjoining them. The subject-predicate pairs are designated here as predications.

The primary concern of this system is with relations inferred between predications. The relations scored as discriminations are said to be "simultaneous" in that they all occur within a single thought unit. The term "inferred" designates a logical analysis of the relational terms. That is, it is an attempt to contrast a logical analysis with an analyses of the semantics alone. While the production of different connectives is scored in terms of different numbers of discriminations inferred from their use, there are no grounds in this model to assert that the speaker is discriminating components which are assessed in the analysis. That is, the system is used as a heuristic device to order, reliably and systematically, inferred differences between utterances, without addressing itself to whether speakers and listeners would themselves report or concur with the discriminations made by the system.

The ten kinds of relations examined and scored in the present

study were: conjoining, temporal ordering, comparison, contrast, causality, teleology, logic, contingency, qualification (i.e., uncertainty) and invariance (each of these is described and exemplified in Table 1). The scale allows comparison of these different kinds of relations, each of which is taken as a separate discrimination, across different specific usages of the "same" word by assigning different scale values depending on the number of relations inferred in each different usage of that word.⁴

Given the definition of a discrimination in this system as an inferred relation between predications, the numerical score of complexity assigned to a thought unit is the total number of inferred discriminations identified in that thought unit. Any lexical item or syntactic construction from which a relation between predications is inferable is scorable in this system. Several disclaimers and qualifications are important to note here.

First, a lexical item per se is not what is scored; rather, an inferred relation is the scorable element. From the use of some lexical items, to be exemplified below, we infer more than one relation between predications.

Second, this system does not attempt to assay any complexity which may be inferred from types of discriminations other than those made about the relations between predications. Thus, the discriminations inferable from all nouns and most adjectives are excluded. Since it excludes such a large segment of discriminations in language use, referring to it as a scale of "language

complexity" may be somewhat misleading. Hopefully, however, the system will allow some control of the component of linguistic complexity with which it does concern itself, especially when this parameter is related to the comprehension of utterances.

Third, all related predications need not occur in the same thought unit for the relations between these predications to be scored. There is an easy opportunity for confusion here. All of the inferred relations which occur in one thought unit are summed to yield a complexity score for that unit; the predications which are being related need not occur in that one unit. A "thought unit" is an arbitrary boundary established to make possible the counting of discriminations within it. In a stream of speech the unit of analysis which has been isolated by definition may contain predications which are related to predications occurring in other thought units. If only the relations between predications occurring within the same thought unit are scored, than a large number of relations across thought units, identical to those scored within thought units, may be omitted from analysis. This omission seems unnecessary and arbitrary. Also, this limitation would result in virtually excluding all discriminations except those inferred from connectives. Connectives have already been subject to an excellent analysis by Bennet (1974), though her principles for ordering are somewhat different.

Fourth, not only are relations inferred by this system but, entire predications in some instances are also inferred; once a predication is inferred, then the relations between the inferred

predication and any fully explicit predication are included in the complexity score. The rationale for inferring predications is similar to the rationale for including relations between predications occurring in different thought units, viz., it permits inclusion of discriminations which can be inferred reliably from particular utterances. A comparison of instances which vary in the degree of explicitness of the related predications will illustrate the usefulness of this procedure.

The most explicit form is one in which both predications (and the connective from which relation between them is inferred) are all explicit: e.g., "He left but she stayed." The next most explicit form is one in which only one predication is fully explicit and the second predication is only partially explicit: e.g., "He tried to reach it but couldn't."

In the following examples, one predication is fully explicit and the second is wholly inferred:

- a) "She remembers where she left her keys."
- b) "She seemed feckless."
- c) "She went back into the closet."

We argue that the following non-negated⁵ predications (which are shown explicitly below) can be inferred reliably from these examples:

- a) "She remembers (i.e., now knows) where her keys are, but for some unspecified period of time prior to now, she did not know where her keys were."⁶
- b) "She seemed (i.e., appeared) feckless, but whether or not she was feckless is uncertain."
- c) "She went back into the closet after she had emerged from the closet at some unspecified earlier time."

Finally, this logic brings us to instances in which both predications are inferred. Consider, "There were these two doll-like figures..." A contrast is made between "figures" and "dolls;" this contrast (one of the scored relations) is given in these two inferred predications: "These figures have some of the characteristics of dolls, but they do not have all of the characteristics of dolls." In this example, two predications are inferred from a modifier and a contrast (the relation between them) is scored.

The sources of inferred predications exemplified above (i.e., particular verbs and modifiers) are elements of a set of lexical items which appear to represent elliptical or condensed predications. We have been unable to locate nominal which permit such inferences. Positing a class of lexical items which are taken to be elliptical predications (e.g., remember) may appear to contradict the earlier statement that lexical items per se are not the scorable discriminations in this system. It must be emphasized that these lexical items are relevant only in that they permit reliable inferences about relations between predications, even though the predications themselves are inferred.

In addition to these lexical items, there is another source of inferred predications, namely, relevant tense variations. In general, tense discriminations are not scored, just as nominal discriminations are not. However, the use of some tenses allows the same reliable inferences of relations as does the use of some connectives. The past perfect progressive allows for the discrimination, in one thought unit and without the use of a connec-

tive, of a temporal ordering of events. (e.g., "She arrived and he had been sitting there" allows the inference, "He had been sitting there before she arrived.") The past progressive allows for the temporal sequencing of events. The simple progressives (past and present) indicate the continuation of an event over time, but do not allow temporal sequencing of the two predications. The atemporal present tense allows the inference of an invariant relation (e.g., "A woman reaps what she sows," is the equivalent of, "If a woman sows something, then that is what she shall reap.") Finally, the subjunctive tense allows for the inference of uncertainty (e.g., "I would like to finish," implies, "I want to finish, but I may not.")

Although the relations between inferred predications are recognized in the complexity score, they are treated somewhat differently than relations between explicit predications. It is assumed that, other things being equal, an explicit statement is more "complex" than an implicit one. Consequently, when the predications are explicit (e.g., "She seemed feckless, but whether she actually was is uncertain,") then the connective (here, but) is scored not only for contrasting the two predications, but also for conjoining them (i.e., the utterance is assigned a score of 2.) When the predications are inferred, the conjoining of the inferred predications is not scored; only the relation is scored (e.g., "She seemed feckless," is assigned a score of 1 for the uncertainty indicated, but is not scored for the conjoining of two predications.)

More generally, it may be noted about this system that the particular set of scorable relations between predications was derived initially from an analysis of English connectives. While it may be that other relations between predications may be reliably inferred, those relations which may be inferred from the use of a connective are of particular interest since the discriminations of these relations may occur simultaneously. Thus, "complexity," as used here, is not merely a function of the absolute number of discriminations (i.e., inferred relations), but is rather a function of the number of inferences which can be said to be made simultaneously. Given the definition of a thought unit, the only way in which two predications can be conjoined within a single thought unit as if they are conjoined by an "appropriate connective other than and." This definition has consequences for the scoring of complexity of a speech sample in which connectives are used sparsely: not only is such a speech sample scored as less complex than one in which connectives are used more frequently, but additionally, the use of few connectives has the effect of shortening the thought unit, resulting in summing scores over a shorter boundary. However, the goal of the scale is to obtain an estimate of the maximum complexity of a given language sample, rather than an average across the sample. It is assumed that other usages at the same level as the maximum complexity usage for a given individual, as well as usages below that maximum level, should be comprehended by that individual. Further, use of an average complexity score would not be as

sensitive to the simultaneity of discriminations. The most "efficient" way of indicating several relations between two predication is by the use of a connective, which indicates simultaneously the relations between predications. It is proposed that the best estimate of the likelihood of comprehension of a given utterance by a given individual is derived from independent assessments of the maximum complexity usage of that individual and the complexity of the language-to-be-comprehended. Predictions of comprehension are made to the extent there is a "match" between these two independent assessments.

One set of concerns of the present study was to test whether (a) this scale could be applied reliably by judges to children's descriptions of different kinds of event; (b) to determine whether children's utterances can be ordered reliably over the different kinds of verbal productions; and (c) to examine the change in complexity from 4th through 8th grades.

Three different kinds of stimulus conditions were used here to maximize the use of relational terms and explore the possibility that the complexity of linguistic description varies as a function of the "complexity" of the stimulus event to be encoded. The three stimulus conditions (viz., animated film, picture frame series, single picture) may be regarded as differentially visually explicit in terms of the relations between events depicted in them. The animated film may be regarded as most explicit in depicting all transitions and sequencing of events (i.e., the visual array is one of the continuous movement). The relations between events may be regarded as less explicitly given in the series of picture

frames. A static visual depiction is presented in each frame, as "slices" of continuous action. The frames depict a "story" in the sense that the scene depicted in the first frame is temporally prior to the scene depicted in the second frame, and so on. Here, the "events" are given explicitly in the stimulus, but the relations between events must be "constructed" (i.e., inferred) by the speaker. The single card task is the least explicit, in that it does not depict a series of events at all; one event is visually presented and the task requirement of relating events is verbally communicated. Here, the speaker must produce events not visually depicted and also related these events to what is visually explicit.

Ordering the stimulus events in terms of visual explicitness does not immediately address itself to which of the stimuli is to be regarded as most complex. In contrast with the ordering of stimuli by analysis of visual explicitness, an examination of the explicitness of task requirements ("Tell a Story") yields a reverse ordering of the stimuli. That is, the single picture has the most explicit instructions for the child, with the motion picture implicitly requiring sequencing, and the picture frame series permitting either sequencing, or single-frame responding, as if independent events.

To explore differential complexity of usage by individuals assumed to differ in language competence, fourth, sixth and eighth graders were asked to make up stories for each of the three kinds of stimuli. It seems tenable to assume that these

groups differ in number and kinds of relations which they encode. Since this scoring system is regarded as a means of assessing such discriminations, it is expected that complexity as assessed with this scale should vary with age (i.e., posited language competence).

The second set of concerns here deals with differences in complexity between children with assumed different socialization backgrounds. The literature is replete (e.g., Bernstein, 1964; Schatzman & Strauss, 1955) with many reports of significant differences in verbal communication contents, contexts and patterns by children from middle versus lower socioeconomic backgrounds. Schatzman & Strauss note, for example, "... the difference is not simply the failure or success -- of lower and upper groups respectively -- in communicating clearly and in sufficient detail for the interviewer's purpose. Nor does the difference merely involve correctness or elaborateness of grammar or use of a more precise or colorful vocabulary. The difference is a considerable disparity in (a) the number of kinds of perspectives utilized in communication; (b) the ability to take the listener's role; (c) the handling of classifications; and (d) the frameworks and stylistic devices which order and implement the communication" (p. 330).

While one may argue that different groups have different concerns in communication and that for one group or another the kinds of categories of concern noted, for example, by Schatzman and Strauss are not important in their own communication contexts,

when children from such backgrounds get to a school where the requirement is "Standard English" usage, then children whose patterns do not include such usages will be at a distinct disadvantage unless the educational system makes it possible for the child to "catch up" or introduces some other compensatory procedures. There is reason to believe that in what has been called "Restricted Code" usage, much complexity may be carried in nominals, rather than in relational terms. When children with such usage come to a school system where "Standard English" is the requirement (and Standard English seems to carry much of its complexity in relational terms), such children may be at a disadvantage. As Wiener and Shilkret (1974) have noted, when there is a mismatch between the complexity (i.e., lesser) of the usages of the comprehender and the material to be comprehended (i.e., greater), non-comprehension will occur. In this investigation, then, a second concern was exploring the differences between two socioeconomic groups in complexity of usage, as assessed by the proposed scale. It is expected that middle and lower class children differ in the degree of complexity evident in their verbal production and that this difference increases (rather than decreases) in the standard school context.

Method

Subjects

Subjects were drawn from two quite distinct sets of neighborhood schools, one from a middle-class socioeconomic background, the other from a lower socioeconomic background. While both

neighborhoods were homogeneous with respect to income levels, there was considerable diversity of ethnic backgrounds, and in this sense neither group of children can be said to come from "ghettos." For each socioeconomic group, 4th and 6th grade children were selected from an elementary school in the immediate neighborhood; the 8th grade children, selected from the middle school, had attended the target elementary school. A total of 135 children participated in the study, divided as follows: (a) middle-class subjects: 20 fourth graders (10 boys; 10 girls); 23 sixth graders (11 boys; 12 girls); 21 eighth graders (10 boys; 11 girls); (b) lower-class subjects: 29 fourth graders (19 boys; 10 girls); 20 sixth graders (10 boys; 10 girls); 22 eighth graders (11 boys; 11 girls).

Stimuli

All children were presented three different kinds of stimuli in a fixed order:

(1) a 6½-minute animated color film without sound track. The film ("Changes, changes", Weston Woods, Inc., Weston, CT) depicts wooden dolls transforming one set of blocks into a variety of objects of their use (e.g., a house, fire engine, train, truck). There are explicit "causes" for each of the transformations (e.g., the house catches on fire, and the dolls move some of the blocks to form a fire engine and direct water at the fire).

(2) A set of seven hand-drawn 3½" X 5½" frames (designed by Joanne Bennet, Clark University) depicting: (a) a small boy, sitting on a box in a playroom, pointing to a shelf above him

which has toys on it; (b) the boy standing with his arms extended to the shelf above; (c) the boy facing the viewer with a disgruntled expression; (d) the boy smiling, carrying a box toward the shelf; (e) the boy standing on one box with his arms extended to the shelf above; (f) a young girl pointing at a second box in the room while the boy remains standing on the first box; (g) the boy, standing on both boxes, smiling, holding the toy which had been on the shelf.

(3) Thematic Apperception Card #1, which depicts a young boy sitting at a table, his head resting on his hands, looking at a violin.

Procedure

Children were tested individually either in a mobile laboratory, an all-purpose room, or in a building near the school. The children were informed that they were to be shown different kinds of pictures and their task was to tell the experimenter what they had seen.

It is to be noted that the primary purpose of this procedure and of the stimuli was to elicit as broad a range of speech as possible in an interaction of 15-30 minutes. To this end, the experimenters attempted to "relax" and engage children in the task by behaving as if child and experimenter were engaged in conversation. That is, experimenters talked in response to the children's verbalizations and this talking was not standardized. In general, experimenters attempted to ask relevant, simple,

open-ended questions. This effort included tactics such as asking questions in ways other than forms which could be answered with a "yes" or "no," or repeating a child's statement in a questioning tone (e.g., "You said his uncle has given him the violin?"). This procedure appeared to be successful to the extent that very few of the children evidenced any difficulty in talking about the stimuli. In that the goal was to assess language complexity, it was not required that the speech samples be limited in any way. Rather, we assumed (a) that the more the child said, the greater was the possibility of producing different speech forms (i.e., the more variable the utterances would be); and (b) as the variability of the utterances increased, the possibility that the child would produce his/her most complex form increased. For the purpose of analyses reported here, only utterances which were about the stimuli were used.

Results and Discussion

Complexity Scoring and Reliability

Verbatim transcriptions of utterances of each subject were made. After careful explanation of the criteria and examples, five practice typescripts were scored by three judges. After discussion and "corrections," each judge was given a set of the middle-class protocols and was required to (a) locate the most complex sentences for each stimulus condition for each child and (b) assign a numerical complexity score derived from the detailed scoring criteria described in Table 1 to the thought units selected. In this system, a score is given whenever a verbal

form indicates a simultaneous discrimination of a relation between predications. Examples of the forms scored were connectives (e.g., until, because), verb forms (e.g., verbs requiring complements or requiring other predications), modifiers (e.g., sort of, probably), negations, and contrasts. A description of the scoring system is provided in Table 1.

Insert Table 1 about here

Despite occasional instances when judges scored different thought units from the same protocol, reliabilities were high ($r_{j1j2} = .86$; $r_{j1j3} = .83$; $r_{j2j3} = .86$, with a range of scores from 0 to 16). Of the total number of pairwise comparisons of judges' scores, over 75% of the judgments made by any two judges differed by one unit or less and over 89% of the judgments by any two judges differed by two units or less. Of the total number of possible pairwise comparisons, approximately 6% differed by more than three units. These were examined individually, in terms of the three judges' responses to each of the protocols. Of this 6% of the total number of pairwise comparisons, it was found that in 39% of the instances only one pairwise comparison (of three differed by more than three units, and in 61% of the instances two pairs of judges differed by more than three units. In no instance was it found that all three pairwise comparisons differed by more than three units. Otherwise stated, a pattern of scoring, for example of $j_1 = 3$; $j_2 = 7$; $j_3 = 11$, never occurred. Rather, one judge was always the "odd man out" in these instances and the

other two judges never disagreed by more than two units.

These judges then trained an additional group of three judges. After training, the six judges scored varying numbers of protocols of all children using the same criteria. For this scoring, if any judge scored a different thought unit of a given protocol, two or more judges also scored that thought unit. Given the relatively high inter-judge reliabilities, and the pattern of disagreements when they did occur, the score assigned a subject for each stimulus condition was the highest complexity score agreed upon by the majority of judges scoring the identical thought unit (i.e., $j_1 = 10$; $j_2 = 11$; $j_3 = 13$ would be scored 11, as the highest score agreed upon by two of the three judges; or $j_1 = 4$; $j_2 = 4$; $j_3 = 6$ would be scored 4).

Analyses of Complexity Scores

The scores thus derived for the middle socioeconomic class and lower socioeconomic class children, for each of the three grades, 4, 6, and 8, for the three tasks were analysed in a 2 (social class) X 2 (sex) X 3 (task) X 3 (grade) analysis of variance with the task factor a repeated measure and the grade factor analyzed with trend components.

The Analysis of Variance showed no significant effects for sex. For the socioeconomic class variable, children from lower-class background evidenced lower complexity (" $\underline{M} = 4.48$ "; than children from middle-class background ($\underline{M} = 5.64$; $\underline{F}_{1,123} = 15.39$, $p < .001$ "). There was a significant linear trend in the mean scores for the three grades ($\underline{M}_4 = 4.57$; $\underline{M}_6 = 5.20$; $\underline{M}_8 = 5.42$; $\underline{F}_{1,123} = 5.51$; $\underline{F}_{1,123} = 5.51$; $p < .025$). Although the social class and

grade interaction did not reach significance ($F_{2,246} = 2.09$; $.10 < p < .20$), an analysis of the means showed a linear increase for middle-class children over grade ($M_{4M} = 4.98$; $M_{6M} = 5.76$; $M_{8M} = 1.18$; difference between M_4 and M_8 significant at $p < .05$ by Duncan's Multiple-Range test). In contrast lower-class children's scores apparently did not increase beyond 6th grade ($M_{4L} = 4.16$; $M_{6L} = 4.65$; $M_{8L} = 4.65$). Incidentally, the means of the 6th and the 8th grade middle-class children were higher than the equivalent grades means for the lower-class children ($p < .05$ or less). Additionally, lower-class children showed less variability than middle-class children.

There was a significant difference for the three tasks ($M_{\text{film}} = 4.35$; $M_{\text{frames}} = 5.64$; $M_{\text{TAT}} = 5.21$; $F_{2,246} = 11.43$; $p < .001$). A Duncan's Multiple-Range procedure indicated that the film elicited less complex productions than both the picture-frame sequence ($df = 246$, $p < .001$) and the TAT card ($df = 246$, $p < .005$).

Correlation coefficients were also computed among the three tasks. The correlations for all subjects were: $r_{\text{film, frames}} = .33$; $p < .01$; $r_{\text{film, TAT}} = .24$; $r_{\text{frames, TAT}} = .17$; $p < .05$; all $df = 133$; two-tailed tests). Coefficients were also computed for each socioeconomic class separately. These results showed that even these relatively small overall correlations were due primarily to the middle-class children's relatively greater consistency across tasks; correlations between tasks for lower-class children were lower and not significant (for middle-class chil-

dren: $r_{\text{film, frame}} = .35$; $p < .01$; $r_{\text{film, TAT}} = .34$; $p < .01$; $r_{\text{frames, TAT}} = .17$; n.s., all $df = 62$; for lower class children: $r_{\text{film, frames}} = .18$; $r_{\text{film, TAT}} = .03$; $r_{\text{frames, TAT}} = .07$; all n.s., $df = 68$).

The interjudge reliability data indicate that the scoring system proposed here meets the first requirement for a measuring instrument -- within a reasonable degree of tolerance, judges agree on a score. Further, the measurements of the complexity of language obtained with this instrument make it possible to differentiate and order groups who are expected to differ in complexity of speech they produce (i.e., across grades). That the grades were found to differ despite the relatively gross criterion for group membership (i.e., age-appropriate grade level) suggests that if other factors had been controlled (e.g., matching for IQ, achievement test scores⁷), differences between group means might have been more marked.

While a difference in mean complexity scores across stimulus conditions was observed, the effects were not ordered in the sequences predicted by the analysis of the visual explicitness of relations depicted in the stimuli. That is, the stimulus which was most visually explicit in this regard (i.e., the film) elicited less complex productions than either of the other two stimuli; and the stimulus of intermediate explicitness elicited the most complex language samples. Since only one member of each stimulus class was used, it is impossible to determine whether these observed differences are particular to the stimuli used or are a

function of class characteristics of the different stimuli. However, examination of the protocols indicated that for the picture frame sequence, the utterance, "He tried to reach the shelf but he couldn't (because he was still too short)," either with or without the causal predication in parentheses, was produced frequently by middle-class subjects across grades. In the system, the shorter utterance is scored 5, the longer utterance 9; both these scores are higher than the mean scores for those subjects for three grades. It may well be that the higher scores in this condition can be accounted for by the content of the frames, rather than the degree of visual explicitness of relations of a frame series as a generic class. A study is in progress which equates as much as possible the content between stimulus conditions (i.e., picture frames are taken from frames of the film).

A second explanation may be offered in terms of more basic differences in the classes of stimuli. While it was originally thought that the film, among the stimulus types, is most visually explicit in depicting relations between events, this very continuity may be said to render "events" depicted in the film as less discrete than events depicted in the picture-frame series. It might be argued that presenting events in discrete frames completes part of the task of encoding for the speaker by obviating the "decision" of what events are to be encoded. Furthermore, the instruction to tell a story may be constructed as a requirement to relate events depicted as isolated by virtue of the fact that they are presented in different frames; for that stimulus,

the events, in a sense, are "given", and the relations are "new". There is the further complication that quantitatively more events were depicted in the film than in the picture frames. Again, an equating, as much as possible, of the "amount" of information between the two stimuli, currently under investigation, would seem to be warranted. It should be remembered, however, that the primary purpose of this study was to articulate a system for assessing complexity of connective terms, and to test the utility of that system in differentiating groups which would be expected to be different in language usage. To that end, a variety of stimulus conditions were used for each subject. To the extent that stimuli are "equated" for content, length, and other variables, they would seem to be less useful in eliciting the variety of language behavior desired in this study (i.e., in a repeated-measures design).

The findings that lower-class children in this study produced less complex utterances than middle-class children should not be taken to imply that lower-class "language" is less "complex" than middle-class "language". Among other reasons, such a conclusion is unwarranted since it would presume that this particular system for assessing complexity is representative of the language in toto. In fact, we would suggest that class differences in language might be fruitfully explored in terms of the differential "loci" for "carrying" the complexity. That is, the complexity (in terms of simultaneous discriminations) of lower-class language may well be carried more in the nominal terms, or

even in intonational contour (see Brooks, Brandt, and Wiener, 1969) and gesture (see Wiener, Devoe, Rubinow, and Geller, 1972) than in the relational terms studied here. The system described here might be expanded to include linguistic discriminations other than those entailed in relating two predications, such as the choice of particular nouns, verbs, and modifiers.

This notwithstanding, the differences obtained in this study between children from different socioeconomic backgrounds are important to consider when a child from one kind of background is placed in an environment where the language of a different background is used. To the extent that middle-class English "carries" complexity relatively more in relational terms than does lower-class language, and the "loci" of complexity of lower-class language are not responded to in a school system, one would expect children from a lower-class background to be at a distinct disadvantage when entering such a system. The evidence from this study suggests that this is the case, and that this disadvantage may become greater as the child progresses through the system.

FOOTNOTES

1. This research was supported by a National Institute of Education Grant NIE-74-0017 to Wiener and Shilkret. Articulation of the scale and test with middle-class children was part of a thesis submitted by Rubano to the Department of Psychology, Clark University, in partial fulfillment of the requirements for the M.A. degree. Grateful acknowledgement is expressed to Gail Hornstein and Gary Yablick who provided criticism of the scale and served as judges for the middle-class data; also to Lynda Godkin, Suzannah Heard, and Laurie Lockwood who served as scorers and assisted with other aspects of the research. Parts of this paper were presented to the 85th Annual Meeting of the American Psychological Association, San Francisco, August 1977.
2. The necessity for this qualification is elaborated in the discussion.
3. There are two problems with this definition of "thought unit." One problem concerns the use of the term "appropriate" in the definition; the other concerns reliability in identifying S-P pairs. The first problem is discussed later. It may be noted here that despite these problems, reliability among judges in identifying thought units is high.
4. For example, if may be construed as "Whether or not," indicating uncertainty; in other instances, if may be construed as indicating a contingent relation, rather than uncertainty (e.g., the if in "He didn't know if he had change in his pocket," relates the two predications and indicates uncer-

tainty; it is scored 2; in contrast, the "same" word, if, in "If the violin is held in the right way, it is easy to play," relates two predications and also indicates a temporal ordering and a contingency; a score of 3. The second example is regarded as more complex than the first.)

5. Inferring a negated predication, e.g., "She is not unaware of the location of her keys," is viewed as a semantic paraphrase (i.e., offers no new information) and is trivial with respect to this analysis. It may be the case that the occurrence of the double negative serves to deny an implicit presupposition (see "negation of a verb form" in Table 1), but this issue is relevant only to explicit and not to inferred predications.
6. The reader may notice that from each of these inferred predications, yet another predication may be inferred. In the first example, the predication explicitly given is, "She remembers...etc." The predication which is inferred is, "She did not know where her keys were before." The use of remember (rather than, e.g., found out) rules out the possibility of some other agent having informed her. Strictly speaking, the predication which should be inferred from "She remembers..." is: "At some previous time she forgot..." Once having inferred, "She forgot..." from the predication explicitly given in the utterance, one is forced by the same logic to infer a third predication, namely, "She must have known at some previous time where her keys were." This

third predication is inferred from the second predication, which is itself inferred from the explicit utterance. To limit the possibility of continuing to infer predications, the arbitrary rule is that the system will infer only one predication. This same argument applies to example C on page 5.

7. Shilkret and Wiener (1972) report significant correlations between reading test and intelligence test scores. It seems reasonable to assume that investigation would show similar correlations between measurements of language complexity and intelligence test scores.

References

- Bennet, J. Usage of verbal connectives by children in relation to age and contextual variation. Unpublished M.A. Thesis, Clark University, 1974.
- Bernstein, B. Elaborated and restricted codes: Their social origins and some consequences. American Anthropologist Special Publication, 1964, 66 (2), 55-69.
- Brooks, R., Brandt, L., & Wiener, M. Differential response to two communication channels: Socioeconomic class difference in response to verbal reinforcers communicated with and without tonal inflection. Child Development, 1969, 40, 453-470.
- Rubano, M. The assessment of linguistic complexity in terms of discrimination of relations. Unpublished M.A. Thesis, Clark University, 1976.
- Schatzman, L. & Strauss, A. Social class and modes of communication. American Journal of Sociology, 1955, 60, 329-338.
- Shilkret, R., & Wiener, M. The contribution of syntactic and para-syntactic cues in the comprehension of spoken and written language. U.S. Department of Health, Education, and Welfare Office of Education Grant No. OEG-1-70-0009 (509), Final Report, February, 1972.
- Wiener, M., Devoe, S., Rubinow, S., & Geller, J. Nonverbal behavior and nonverbal communication. Psychological Review, 1972, 79, 185-214.
- Wiener, M., & Shilkret, R. Complexity in auditory and graphic language. National Institute of Education Grant Proposal, No. NIE-G-74-0017, 1974.

Table 1

Scoring System for Language Complexity

Category	Rationales/Examples	Score
----------	---------------------	-------

Connectives

The following is a partial list of connectives and the number of discriminations entailed in the use of each (i.e., the score to be assigned to each).

<u>and</u> <u>and then</u>	These forms are the simplest and most ubiquitous connective. For purposes of these data they are not to be scored since they do not lead to a differentiation of Ss. (Further, some occurrences of <u>and then</u> , or <u>so then</u> , appear susceptible to categorization as narrative devices or gap fillers with the result that they artificially conjoin clauses).	0
<u>and not</u>	Serves to relate and contrast two predications. Treated as a connective on the order of <u>but</u> , but less specific. The second clause is related to the first as the negation of (an implicitly) predicted outcome. E.g., "He tried to reach the truck <u>and</u> he couldn't." (Note: <u>And not</u> , while treated as a connective, signals the beginning of a new thought unit. Also it is not to be confused with the case in which a non-negated predication and a negated predication are simply linked with <u>and</u> , e.g., "They build a truck to carry them <u>and</u> they couldn't go farther, so they stopped.")	2
<u>after</u> <u>before</u> <u>when</u> <u>since</u> <u>while</u> , <u>as</u> , (when used tem- porally)	Each serves to relate two predications and specifies a temporal relation. " <u>After</u> his mother called his, he went in." E.g., " <u>As</u> he was waxing the car, it began to rain."	2
<u>as if</u> <u>like</u>	Relation; comparison. E.g., "They looked <u>like</u> they liked each other." (Note: <u>Like</u> is not scored when used as a pause filler or narrative device, e.g., "It was <u>like</u> , you know, well <u>like</u> I think they liked each other.")	2

Table 1 (continued)

Category	Rationales/Examples	Score
Connectives (continued)		
<u>as little</u> <u>as</u> <u>as much as</u>	Relation; comparative ordering. E.g., "I practice violin <u>as much as</u> I want to."	2
<u>because</u>	Relation; nomic connection of either a temporal or logical type. (Temporal lawfulness is the assertion of invariant temporal co-occurrences.) E.g., "The fourth plug didn't fire <u>because</u> the ignition wire was disconnected." (Temporal) "The article was not stimulating <u>because</u> it contained so many platitudes." (Logical)	3
<u>but</u>	There are four possible scorable predications entailed in any non-anomalous use of <u>but</u> . It is always scored for signaling a relation between two predications; additionally, it is scored as signaling a negation of, or contrast with, the first of the two predications it relates. Consequently, the use of <u>but</u> entails minimally two predications.	
<u>but</u> (con- tempora- neous)	Relation; contrast (where the contrasted item is not given in the same utterance.) E.g., (Did you eat dinner with Jack?) "I indulged, <u>but</u> he abstained."	2
<u>but</u>	Relation; contrast; temporal sequencing (where the semantics of the clauses indicate a temporal priority.) E.g., "I went to the bank to make a deposit, <u>but</u> it was closed." (From which it can be inferred that the bank closed <u>before</u> I got there.)	3
<u>but</u>	Relation; nomic connection (of an unspecified type); negation of the nomic connection (i.e., contrast with the law implied.) E.g., "The ball was round, <u>but</u> it didn't roll."	3
<u>but</u>	Relation; nomic connection, negated; temporal sequencing. E.g., "He failed his exam, <u>but</u> he had studied." (The law being negated here is, "If one studies for an exam, one [subsequently] passes it.")	4
<u>in order</u> <u>to</u> <u>to</u> <u>so that</u>	Relation; nomic connection, teleological. E.g., "He reached up <u>to</u> get the truck." "He stood on the box <u>so that</u> he could reach the truck."	3

Table 1 (continued)

Category	Rationales/Examples	Score
Connectives (continued)		
<u>if</u>	There are four possible scorable predications entailed in any non-anomalous use of <u>if</u> . It is always scored for signaling a relation between two predications; additionally, it is always scored for indicating either uncertainty or contingency. Consequently, the non-anomalous use of <u>if</u> entails <u>minimally</u> two predications.	
<u>if</u>	(When used to mean <u>whether or not</u>) Relation; statement of uncertainty. E.g., "He didn't know <u>if</u> he had change in his pocket."	2
<u>if</u>	Relation; temporal statement; contingency. E.g., "He finally decided that it just might play better <u>if</u> he turned it backwards." (Note: that the instrument play better is contingent on its being turned backward; that this desired outcome is <u>uncertain</u> is NOT indicated by the <u>if</u> , but by the <u>might play</u> , which is scored independently for uncertainty.)	3
<u>if</u>	Relation; temporal ordering; contingency; nomic connection. E.g., "If the violin is held in the proper position, (<u>then</u>) it is easy to play."	4
<u>so</u>	Relation; nomic connection, logical or temporal consequent. The use of <u>so</u> always implies a law-like connection between the two events; the law may be one that states temporal sequencing for events or it may state logical consequences. E.g., (Temporal) "Their house caught fire, <u>so</u> they made a fire engine." (In this sense, <u>so</u> is the equivalent of <u>because</u> , viz., "They made a fire engine <u>because</u> <u>their house caught fire</u> .") <u>So</u> is also used to introduce logical, atemporal consequences. E.g., "He had never been a member of the armed forces <u>so</u> he was not entitled to Veterans' benefits." (Notes: <u>So</u> is not to be confused with <u>so then</u> , which is always taken to indicate only temporal sequencing or with <u>and so</u> , when either of these connectives is used as a variant of <u>and then</u> . <u>And so</u> appears to be a transitional form between <u>so</u> (consequently) and <u>and then</u> . The following are guidelines for identifying lower complexity <u>and so</u> 's: <u>and so</u> 's are scored as <u>and then</u> 's if:	3

Table 1 (continued)

Category	Rationales/Examples	Score
Connectives (continued)		
	(1) no other <u>so</u> appears, used appropriately in the protocol. (2) sentences (indicated by periods in transcription) BEGIN with <u>and so</u> OR <u>so</u>, even if the <u>and so</u> at the beginning of the next utterance is a proper consequent of the previous utterance. (3) Even if neither of the above two conditions is fulfilled, <u>and so</u> may be scored as <u>and then</u> if it appears in a 'run-on' sentence as determined by the judgment of the scorer. E.g., "I don't know what happened to the car but then they turned it into a train and they got to the end of the track and there was no more track and <u>so</u> they built themselves a house again." This <u>and so</u>, even though it introduces the consequent of running out of track, is regarded as a 'run-on' and treated as an <u>and then</u>.)	
<u>until</u>	There are four possible scorable predications entailed in any non-anomalous use of <u>until</u> . It is always scored for signaling a relation between two predications; additionally, it is scored for at least one of two types of temporal reference. Consequently, the non-anomalous use of <u>until</u> entails minimally two relations.	
<u>until</u>	(When there is no specifiable temporal contingency.) Relation; temporal marking (as opposed to temporal contingency.) E.g., "I didn't get to school <u>until</u> the colloquium was over."	2
<u>until</u>	(When there is a specifiable temporal contingency marking an instantaneous point in time.) Relation; temporal ordering; temporal contingency. E.g., "He didn't open the door <u>until</u> she left." (In this instance, the ordering of events is given, i.e., she left and then he opened the door, as well as the fact that his opening the door was contingent on her leaving.)	3
<u>until</u>	Relation; temporal ordering of events; continuation of an action over time; temporal contingency. E.g., "He played <u>until</u> he knew his mother wasn't listening anymore." (In addition to stating the order of events, i.e., his mother stopped listening and then	4

Table 1 (continued)

Category	Rationales/Examples	Score
----------	---------------------	-------

Connectives (continued)

<u>until</u>	he stopped playing, and in addition to stating that one event continued over time, i.e., his playing, this <u>until</u> also indicates that the occurrence of one event, his mother not listening, was the condition under which another event transpired, i.e., he stopped. That is, this use of <u>until</u> relates two events, one of which is given as a precondition of the other.)	
--------------	---	--

<u>what, who</u> <u>which</u> <u>how, that</u> <u>where</u>	Each serves to relate two predications. E.g., "he found out <u>how</u> it was done." (Note: When these words are used to introduce relative adjectival clauses they do not serve to relate two complete predications and therefore are not scored. E.g., "It was about two dolls <u>who</u> made a house." Also NOTE: <u>That</u> is optionally deleted between English clauses. (He said <u>that</u> he was coming. He said <u> </u> he was coming.) Sentences in which it is deleted are scored nevertheless because of the relation between the clauses.)	1
--	---	---

Verb Forms

<u>learned to</u> <u>learned</u> <u>how to</u> <u>know how</u> <u>to</u> <u>decide to</u>	These verbs belong to the class of verbs requiring verbal complements. The verb and the complement are taken as the two predications which are related. E.g., "His father <u>let</u> him <u>play</u> with his friends."	1
--	---	---

<u>help to</u> <u>require to</u> <u>tell to</u> <u>let (permit to)</u>		
---	--	--

<u>remember</u> <u>become</u> <u>remake</u> <u>make into</u> <u>get (become)</u> <u>grow</u>	These verbs belong to the class of verbs which imply other predications. They either indicate a state of transition or predicate a change of state such that a non-negated predication about some previous state can be inferred. E.g., the statement, "They turned the truck <u>back</u> into a house," permits the statement, "The truck was previously a house." (Contrast this with, "He ran to the store." The only inferences permissible here are negated ones, e.g., "He was <u>not</u>	1
---	---	---

Table 1 (continued)

Category	Rationales/Examples	Score
Verb forms (continued)		
	at the store when he began running." "He did <u>not</u> walk to the store.")	
<u>want to</u>	In addition to belonging to the class of <u>verbs requiring verbal complements</u> , these verbs may also <u>imply other predications</u> . Namely, <u>want to</u> and <u>try to</u> may imply that the complement has not transpired. E.g., "He <u>wanted to</u> reach his truck and couldn't." However, they do not <u>invariably</u> imply that the complement has not transpired: e.g., "He tried to learn a more difficult piece and succeeded quickly."	1
<u>seemed to</u> (uncertainty) <u>supposed to</u> (moral law) <u>have to</u> (necessity) <u>can</u> (potentiality)	Besides belonging to the class of <u>verbs requiring verbal complements</u> , these verbs also <u>invariably imply other predications</u> . For example, <u>can</u> not only requires a verbal complement (can do something) but invariably makes an assertion about the capability of an actor, the possibility of an event or some sufficient attribute of an object.	2
present tense	Atemporal present tense is a <u>relevant tense variation</u> and is scored for <u>nomc assertion</u> when it indicates a general case statement. E.g., "A man reaps what he sows." (Note: this is not to be confused with a narrative present tense: e.g., (Tell me about the film) "Well, there's a man, and he gets up and looks over at the woman...")	1
past perfect progressive	The past perfect progressive tense is a <u>relevant tense variation</u> (in contrast to the <u>simple past</u> or the <u>past progressive</u>) in that it allows the inference that the event referred to occurred repeatedly or continuously over time in the past. E.g., "He <u>had been sitting</u> there." "He <u>used to go</u> over and borrow his friend's violin." (Note: the general principle is scoring progressives is that they allow the inference of repetition or continuity. The present progressive and the simple past progressive-- "He was sitting there" -- do not imply continuation but rather indicate a specious non-enduring present or a particular point in past time.)	1

Table 1 (continued)

Category	Rationales/Examples	Score
----------	---------------------	-------

Verb forms (continued)

subjunctive tense	The subjunctive tense, although now defective in English, is a <u>relevant tense variation</u>. The subjunctive forms indicate a hypothetical or uncertain quality of an event, both of which are scorable inferences (implied predications) in this system. <u>Could</u> and <u>should</u>, for instance, WHEN THESE FORMS ARE USED TO MARK THE SUBJUNCTIVE TENSE, are scored 1, in addition to other scorable items (e.g., <u>requiring a verbal complement</u>, <u>implying other predications</u> about conformity to a standard, potentially, etc.) (NOTE: care should be taken not to confuse identical forms. For instance, <u>would</u> marks both the subjunctive of <u>will</u> and the past progressive tense, e.g., "He'd (he would) get up every morning at 6:00 in the summer to jog. In the winter, he got up at 7:30" <u>Could</u> is scored for uncertainty where the three following conditions obtain: (1) <u>could</u> is used as the past tense of <u>can</u> (2) the predication is not negated (3) appending, "...but didn't," yields a non-anomalous utterance. E.g., "Given the time he left here, he <u>could have</u> arrived there by 7:00."	1
-------------------	---	---

Modifiers, adjectives and adverbs

<u>kind of</u> <u>sort of</u> <u>like</u> <u>as much as</u>	The use of these words implies a discrimination between the object and the class of objects to which it is compared, i.e., the object is both similar to and different from other members of the class.	1
<u>other</u> , <u>own</u>	Modifiers in general are <u>not</u> scored. Scorable items are to be distinguished from other adjectives and adverbs in that they have clear reference to some previous predication. E.g., "His parents didn't want him to play violin but when he grew up, he went his <u>own</u> way."	1

Table 1 (continued)

Category	Rationale/Examples	Score
Modifiers, adjectives and adverbs (continued)		
<u>probably</u> <u>perhaps</u> <u>possibly</u> <u>conceivably</u>	All these adverbs convey <u>uncertainty</u> about the predication they modify and are scored for that reason.	1
<u>still</u> <u>for a</u> <u>while</u> <u>again</u> <u>all of a</u> <u>sudden</u>	Temporal markers are scored on the same principles enuciated for other scorable temporal items (i.e., the tense variations and temporal connectives). They may indicate <u>continuity</u> (an action enduring over time) e.g., "He sat down <u>for a while</u> after moving the furniture," <u>contrast</u> , e.g., "He got up on the box and he <u>still</u> couldn't reach it," or <u>repetition</u> , e.g., "He <u>tried</u> to reach it <u>again</u> after getting on the box."	1
<u>to think</u> <u>that</u> <u>to suppose</u> <u>that</u> <u>to believe</u> <u>that</u>	Scored for <u>uncertainty</u> on the part of the speaker in addition to other scorable items.	1
negation of a verb form	In general, negation is <u>not</u> scored. However, when the negation has reference to some previous predication (which usually occurs in the denial of an implicit presupposition), it is scored. E.g., "His mother said, 'Very good, you learned to play the violin all by yourself.' 'Well, it wasn't too easy, but I did it.'" The implied supposition which is denied is, "Given you learned to play the violin all by yourself, it must be easy."	1
contrasts	Contrasts of actions, object qualities, nominals are scored even though they do not relate entire predications. E.g., "He tried to reach <u>but</u> couldn't." "It was metal, but lusterless." "I want this pail, <u>not</u> that one."	1

Table 1 (continued)

General Notes on Scoring

- 1) Both clauses must appear for a connective to be scored. In a statement beginning with, "Because..." in response to an inquiry, the because is not scored as a connective since the two predications it relates have not both been produced by the subject. Similarly, when a statement begins with "So..." the so is not scored as a connective if the two predications were not produced as part of the same thought unit.
- 2) Negations following contrasts are redundant. For example, in, "He wanted to reach the shelf but he couldn't," the use of the but as the connective requires the negation in the following clause (i.e., "He wanted to reach the shelf but he could," is an anomalous sentence) and is therefore redundant.
- 3) Quotations will be scored only with respect to relations in the quotations or between them (when they occur in the same utterance) and not on their embeddedness per se. For example, "He sat there thinking, 'What am I going to do? because my mother will say, 'Why aren't you practicing?'" This sentence is reduced, for purposes of scoring, to "He sat there thinking (something because my mother will say (something)), " and would be assigned a score of 3. Direct quotations are taken to be subordinated, but they are not subordinated by the use of a connective; rather, they are subordinated by the very fact of being direct quotations. Therefore, they are considered part of the thought unit, but are not scored for the subordination inherent in their being quotations.
- 4) Where anacolutha (lack of grammatical sequence or coherence) appear, only that part of the utterance which is complete is scored. E.g., "He wants to reach the shelf but he can't find...he's too small." Score, "...but he's too small," and omit, "he can't find..."
- 5) In instances where appositives (juxtaposed phrases) appear in an utterance they are regarded as independent thought units (though they are embedded) and are scored independently. For example, "His sister came, at least I think it was his sister came, and she pointed out that..." The underlined clause in apposition is scored as an independent thought unit in this utterance.

Language complexity of three age groups of middle class children
for two tasks with similar content

Robert Shilkret and Morton Wiener
Clark University

In a previous study (Rubano, Wiener and Shilkret) comparing complexity of language production of children from two socio-economic class groups across three grades, three kinds of stimuli were used to elicit language samples: a short film, a series of picture frames, and a TAT card. Given our principle of taking the most complex utterance as a measure of language "competence," different stimuli were used in the hope of sampling a range of language behavior under different stimulus conditions. In these three tasks, it can be noted that two independent principles characterize differences in task requirements for the subject. On the one hand, the three tasks differ in the degree of explicitness of temporal-sequential occurrences available to the subject. In this dimension, the motion picture is most explicit (e.g., a continuous series of events in time), the cartoon frames are next, but require some construction or "integration" over frames, with the single picture requiring a temporal-logical sequencing constructed with limited information given. On the other hand, for the verbal representation the single picture has the most explicit instructions for the child, with the motion picture implicitly requiring sequencing, and the cartoon frames permitting either sequencing or single-frame responding, as if independent events.

In the original study, while a difference in complexity scores across stimulus conditions was observed, the findings

were not ordered in any way corresponding with the analysis of the visual explicitness of the relations depicted in the stimuli nor of the requirements for verbal representation. That is, the stimulus which was most visually explicit in this regard (i.e., the film) elicited less complex productions than either of the other two stimuli; and the stimulus of intermediate explicitness elicited the most complex language samples (i.e., the cartoon series). Since only one member of each stimulus class was used, it is impossible to determine whether these observed differences are particular to the stimuli used or are a function of class characteristics of the different stimuli. Further, the content or theme of each of the three stimuli used were different, since, in that study, each subject was presented with all three stimulus classes. There was some evidence to suggest that the higher scores for the picture frame series might have been due to the specific content of that stimulus (i.e., utterances like, "He tried to reach the shelf but he couldn't because he was too short" resulted in higher complexity scores for this stimulus).

A second explanation was suggested in terms of more basic differences in the classes of stimuli. While it was originally thought that the film, among the different types of stimuli, is most explicit in depicting relations between events, this very continuity may be said to render "events" depicted in the film as less discrete than events depicted in the picture-frame series. It might be argued that presenting events in discrete frames completes part of the task of encoding for the speaker by obviating the "decision" of which events are to be encoded. Furthermore,

the instruction to tell a story may be constructed as a requirement to relate events depicted as isolated by virtue of the fact that they are presented in different frames, for the picture frames, the events, in a sense, are "given," and the relations are required to be imposed, relations which are the very core of our scoring. There is the further complication that quantitatively more events were depicted in the film than in the picture frames. Again, some equating, as much as possible, of the "amount" of information between the two stimuli would seem to be warranted. It should be remembered, however, that the primary purpose of the original study was to articulate a system for assessing complexity of connective terms, and to test the utility of that system in differentiating groups which would be expected to be different in their language usage. To that end, a variety of stimulus conditions were used for each subject. To the extent that stimuli are "equated" for content, length, and other variables, they would seem to be less useful in eliciting the variety of language behavior desired in assessing individual differences and/or getting a representative sample of a subject's language usages.

Given these considerations, this study was conducted to examine differences in complexity of connective terms of language over age under different stimulus conditions, with "content" and "length" of the stimulus equated as much as possible. To equate content and length as much as possible, a film and a picture-frame series composed of scenes from the film were employed. Middle-class children from the same three age groups as the original study were employed, since results from the previous study

suggested that children from this socioeconomic class background show an increase average in their complexity scores.

Method

Subjects

Subjects were drawn from the school system of an intermediate-size New England town; for the most part students in these schools live in middle-class neighborhoods. Fourth and 6th graders were recruited from two elementary schools; 8th graders were recruited from a third school, an intermediate school, which had students who had come from the two elementary schools used in this study. A total of 106 children participated, divided as follows: 35 4th graders (18 girls; 17 boys), 38 6th graders (14 girls; 24 boys), and 33 8th graders (16 girls; 17 boys).

Stimuli

Two groups of subjects within each age group were established. A child was presented with one of the following stimuli:

(1) A 3½ minutes version of the color film "Changes, changes" (Weston Woods, Inc., Weston CT), the same film used in the first study. The film depicts wooden dolls transforming one set of blocks into a variety of objects for their use (e.g., house, fire engine, boat). The film segment used ended after the "boat" event of the sequence so that a picture-frame series that was not overly long could be presented to a second group of children.

(2) A set of 13 color 9X7½" picture frames, taken from frames of the film and produced by the same company. The pic-

tures depicted:

1. A man and woman wooden dolls standing to the right of various sized, arranged blocks.
2. A man and a woman building a foundation with the blocks.
3. Man and Woman continuing to build house-like structure.
4. Man and woman standing inside their completed house.
5. Smoke and flames breaking out to the left upper side of house.
6. Spreading of the fire; man taking down some of blocks with woman helping.
7. Further spread of fire; man and woman continuing to rearrange some of the blocks.
8. A large fire; man and woman constructing a fire truck with hose from the blocks which made the house.
9. Man putting out fire with hose and water; woman sitting on fire truck.
10. Continuing to put out fire; water level rising.
11. Water at flood level; man and woman rearranging blocks.
12. Man and woman rearranging blocks into a boat.
13. Woman standing on left, man on right, both on boat and waving as boat sails away.

Procedure

Students, all of whom had returned permission slips from parents to participate, were tested individually in all-purpose rooms in their schools. An attempt was made to make the procedure as similar to that used in the Rubano, Wiener and Shilkret study, with the exception that in this study each child saw only

one stimulus. Children in the picture-frame condition were shown the pictures one at a time; this procedure differed from the picture-frame condition of the original study, where all frames were presented at once, with the child instructed to look from left to right. In the present study, a considerably larger number of frames constituted this condition, and the procedure of frame-by-frame presentation was used to better approximate the requirements of the film condition (e.g., memory requirements).

After seeing one of the stimuli, the child was asked to tell what he or she had seen; the child was occasionally prompted to tell more, but in a way that did not add to the "content" of the child's production. After the description, the child was asked whether s(he) liked the film or pictures, and whether s(he) thought it would be good to show third graders. The experimental session, which lasted 10-15 minutes, was taped for later transcription.

Results and Discussion

Complexity Scoring and Reliability

Verbatim transcriptions of utterances of each subject were made. Copies of these transcripts were then scored for complexity by three independent judges who had been trained for this purpose, using the manual and scoring principles derived from the original study. As before, each judge assigned scores to each subject for the utterance considered most complex for that subject. For this study, two sets of complexity scores were assigned for each subject: first, a score for the Description of the stimulus, and,

second, a score for all the subjects' utterances, including the responses to the evaluation questions. For 25 of the 106 subjects, this Overall complexity score was higher than the Descriptive complexity score, and consequently independent sets of analyses were conducted on the two sets of scores.

The scoring procedures includes: (1) independent assessment of complexity by judges, and (2) a discussion to resolve any discrepancies in scoring. As in the original study, discrepancies of more than three or four units were relatively infrequent; the majority of discrepancies which occurred were due to different judges scoring different utterance lengths. A number of discrepancies were due to relatively minor problems (e.g., misreading a word).

Despite such discrepancies, reliabilities were high. For this study, reliability coefficients were computed separately for judges' original scoring (i.e., before discussion and resolution of discrepancies) and for final scoring, and separately for Description scores (range: 0-11) and Overall scores (range: 0-17). Interjudge reliabilities for the original (before resolution) scores were: (1) Description scores: $J_1J_2 = .81$; $J_1J_3 = .74$; $J_2J_3 = .80$; (2) Overall scores: $J_1J_2 = .85$; $J_1J_3 = .71$; $J_2J_3 = .75$. Interjudge reliabilities for scores after discussion and resolution were: (1) Description scores: $J_1J_2 = .97$; $J_1J_3 = .81$; $J_2J_3 = .84$; (2) Overall scores: $J_1J_2 = .97$; $J_1J_3 = .76$; $J_2J_3 = .78$. Thus, while as expected, there was some increase in reliability after discussion, the reliabilities before any communication between judges were relatively high.

Separate sets of reliability coefficients were also computed for each of the three grades; no differences in interjudge reliabilities between grades were apparent.

The complexity score assigned to each child was the highest score agreed upon by at least two judges; employing this principle, two such scores were derived for each child (i.e., Description; Overall). Since the purpose of this scoring procedure was to assay the maximal complexity used by the child, Overall complexity scores included all of the child's language during the session and differed from the Description scores only when some utterance in response to the evaluation questions were higher than the child's Description complexity score.

For this study, in addition to the complexity scores, the language productions were examined in terms of more widely-used measures, sentence length and mean utterance length. Using procedures for scoring "word" and "utterance" outlined by Brown (1973), the length (number of words) of the utterance judges agreed as most complex was derived; the mean length of utterance was computed for each subject, an average over all utterances. For Descriptions of the stimuli, correlation of complexity scores with length of the scored utterance was .72; considering grades separately, these correlations were: 4th grade = .76; 6th grade = .80; 8th grade = .66. For Overall complexity scores, correlation of complexity with length of those utterances was .73 (4th grade = .75; 6th grade = .71; 8th grade = .70). The correlation of Description complexity scores with mean length of utterance was .60 (4th grade = .56; 6th grade = .55; 8th grade = .60).

Finally, the correlation of Overall complexity scores with mean length of utterance was .53 (4th grade = .53; 6th grade = .52; 8th grade = .53). Thus, it would seem that complexity and length are somewhat independent measures, as the concept of complexity used here would imply; independent analyses were conducted on the length measures.

Analyses of Complexity and Length Scores

Each of the response measures was analyzed in an analysis of variance, with the factors of grade (3 levels) task (2 factors pictures vs. film) and sex (male and female), with each variable as an independent groups factor, and trend components for the grade factor. On none of these analyses was task (i.e., pictures vs. film) significant as a main effect or significant in any interaction. Thus, it would seem that equating, as much as was possible, the content of the two kinds of tasks eliminated the task effects apparent in the original study.

For the Description complexity scores, grade was significant but only at the $p < .10$ level, $F(2, 94) = 3.05$. There was a significant quadratic trend component, $F(1, 94) = 11.74$, $p < .001$ (means: 4th grade = 3.61; 6th grade = 5.00; 8th grade = 4.35). It should be noted that the scores for this group of middle class children was somewhat lower than the scores of middle class children in the original study; further, there was not an orderly increase in complexity scores with grade.

For the Overall complexity scores, grade was significant, $F(2, 94) = 3.32$, $p < .05$, with linear trend, $F(1, 94) = 2.78$,

$p < .10$ and quadratic trend, $F(1, 94) = 6.21$, $p < .025$ (means: 4th grade = 4.12; 6th grade = 5.68; 8th grade = 5.62). The interaction between grade and sex was significant, but again only at $p < .10$, $F(2, 94) = 2.77$. A Duncan's Multiple-Range procedure indicated that the girls' Overall complexity scores increased systematically over the three grades (4th grade = 3.89; 6th grade = 5.79; 8th grade = 6.88), while the boys' scores showed no such systematic increase (4th grade = 4.41; 6th grade = 5.57; 8th grade = 4.36).

Examining the length of the most complex descriptive utterance, grade was significant as a main effect, $F(2, 94) = 3.56$, $p < .05$, with linear trend, $F(1, 94) = 3.73$, $p < .10$, and quadratic trend, $F(1, 94) = 5.02$, $p < .05$ (means: 4th grade = 12.05; 6th grade = 14.97; 8th grade = 15.89). No other main effects or interactions were significant.

Examining the length of the Overall most complex utterance (i.e., including responses to the evaluation questions), grade was significant again, $F(2, 94) = 4.60$, $p < .05$, with linear trend, $F(1, 94) = 5.41$, $p < .025$, and quadratic trend, $F(1, 94) = 4.30$, $p < .05$ (means: 4th grade = 12.08; 6th grade = 15.22; 8th grade = 16.76). No other main effect or interaction was significant in this interaction.

Finally, an analysis of mean length of utterance yielded no significant main effects or interactions, although the means across grades "increases" (4th grade = 8.58; 6th grade = 9.14; 8th grade = 9.31).

In summary of these findings, it can be noted that, as in the original study, there was a greater increase in complexity, across all subjects, from the 4th to the 6th grade than from the 6th to the 8th grade. The apparent decrease in complexity among the 8th graders is somewhat puzzling. This decrease was due primarily to the group of 8th grade boys in this study; girls' scores manifest a systematic increase across grades. This pattern of results was particularly apparent in the Overall complexity scores; in effect, 8th grade girls produced more complex utterances in response to questions than they did in a free description of the stimuli, whereas boys' scores showed no such increase. The experimenters for this study were female, and it may well be that the adolescent girls (i.e., the 8th graders) in this study were more verbal, or at least verbally complex, in interaction with female adults than were adolescent boys. It should be noted that the apparent decrease in the 8th grade subjects' complexity scores was not statistically reliable; that is, Duncan's Multiple-Range procedures indicated no significant differences between 6th and 8th graders' scores. Thus, it may be that, given the shorter film used in this study and the generally lower complexity scores, as compared with the original study, a ceiling had been reached for these tasks by the 6th grade. This would suggest the importance of designing tasks and scoring procedures that "scale out" across the range of subjects in subsequent work.

Again, the primary purpose of this study was to study the effects of "equating" task content when assessing the complexity

of language. Results over several response measures in this study indicated that to the extent that content and length of two tasks (producing stories to a film vs. picture series) are equivalent, the complexity of language is equivalent. These results would seem to argue against the analysis of the two tasks discussed earlier. However, it should be noted that the stimuli used in this study were relatively "straightforward" (e.g., nonambiguous; able to be described easily). It may well be that with stimuli that are less clear or which require particular connectives for adequate description (e.g., causal necessity), more complex utterances would be produced. Results of this study would seem to suggest that "content" of the stimuli themselves would be more important in producing differences in complexity than differences in the tasks above (i.e., film vs. picture series).

References

Brown, R. A first language: The early stages. Cambridge, MA:
Harvard University Press, 1973.

Comprehension Difficulty of Three Levels of Text Complexity for
Readers who Differ in the Complexity of their Verbal Production

Rick Schnable and Morton Wiener
Clark University

and Robert Shilkret
Mt. Holyoke College

This study explores a model which attempts to account for comprehension in terms of a match-mismatch analysis of the requirements of the language task, the explicit "language" form available in the text, and the response patterns of the comprehender--all assessed independently. In brief, the framework to be explored examines the failures or successes of comprehension (both graphic and auditory material) less in terms of a deficiency model than in terms of a difference in the pattern the individual (or group) brings to a particular communication instance; that is, the model proposes that comprehension requires a match between the patterns of communication an individual brings to the communication instance and the complexity of the communication to be comprehended.

Within a language community, the behaviors which are taken as "comprehension" seem to be a commonplace occurrence for most individuals and for most communication instances. One rarely considers comprehension as a problem except for children or for others who are usually thought of as defective or deficient. Indeed, if language comprehension occurred for all adults in all instances, one might question whether comprehension would arise as a general problem for psychology rather than as a problem for pedagogy or for individual differences in the sense of differences in "intelligence." That is, the general problem "Under what conditions is comprehension possible?" becomes apparent only in those instances when an adult assumed to be efficient in language usage "miscomprehends" or fails to comprehend a given communication (i.e., and the failure

cannot be attributed to the comprehender as if he were "defective" in some way). However, if one considers this issue carefully, it is clear that almost all members of a language group have difficulty in some kinds of language instances (e.g., a psychologist reading a physics book).

In this context, it is evident that any consideration of language behavior will indicate that a given instance can be characterized by some statement of its "complexity" (e.g., semantic components, syntactic form; sentence length) and the "difficulty" for a given individual (or group) relative to other language instances. It would seem possible, in principle at least, to evolve a scale to compare and contrast all instances of communication in terms of complexity and predictive difficulty for any specific individual. Were such scaling possible, it would follow that for any individual what is taken as comprehension behavior would be seen to occur only for some set of instances; that is, for those instances which represent only some portion of the total possible range of any such scale.

Despite the possibility of comparing any language instance with other language instances and thereby deriving a continuum of "complexity" and apparent predictive difficulty, language users typically seem to respond in an "either-or" manner. If, as we hold, a scalar notion of language usages is meaningful over the whole range of instances despite the "either-or" comprehension of language users, then it would seem necessary to posit some characteristics of a language event (e.g., the intonational contour in speech) which, when present facilitate the possibility of comprehension, and in this sense "compensate" for an increase in the complexity of the forms of the language. Given (1) a conceptualization that language forms can be ordered, and (2) an analysis which makes it possible to denote features of the stimulus which may "compensate" for the increased complexity, it should then be possible to analyze and predict instances

of comprehension and non-comprehension--both for cases when compensatory features are "absent" as well as for cases when these features are "present", as long as we also have some independent indicator of the range of complexity a given individual (or group) has mastered.

Much of the work within the broad field of language comprehension can be seen as derived from one or the other of two quite different assumption models. (1) on the one hand, much recent work has focused on what might be called "characteristics of language" exemplified by the "psycholinguistic" approach. The general procedure of this approach seems to be to select a group of subjects, assumed to be competent and equivalent language users, who are assessed in experimental tasks requiring them to perform with a variety of language instances, ordered by some concept of complexity. Since the subjects are assumed to be identical, performance variability is assumed to derive from differences in the stimuli alone (i.e., between language instances). (2) On the other hand, other work within this field has focused on differences between language users, exemplified by much of the work in the area of reading difficulties, and more generally the recent "sociolinguistic" approach. Often the assumption within this kind of approach appears to be that variability of language performance is to be attributed to individual differences alone (e.g., the characterization of an individual as a "poor reader" with no specification of the material which is or is not comprehended by that individual). Certainly each of these two approaches is valuable; however, adopting either one or the other of them seems to lead to logical or empirical difficulties which are often not apparent until the assumptions are made explicit. For example, recent work which has attempted to adopt a "psycholinguistic" approach to

"reading problems" often also includes a concern with assessing a range of variability both within and between individuals. More often in these kinds of approaches the concern appears to be only with one particular group (i.e., those labelled "poor readers"), without apparent recognition by the investigator that the same kinds of concerns can be applied to within individual instances of comprehension or non-comprehension.

While the focus of our point of view is on the relationship between the general class "event-to-be-comprehended" and the general class "comprehender's usage," we assume that each may be assessed independently, at least from the point of view of an ideal observer. We also assume the possibility for positing an "ideal" scale or ordering the complexity of a language instance in terms of logical relations in the context of the language form. The construction of such a scale assumes that event A' is taken to be more "complex" than event A, if A' includes the discrimination, combination or relation. For example, if an event is a correlation or coordination of two components (X and Y), this instance is taken to be less "complex" by such a scale than, for example, the conditionality of the same two components (X if Y); conditionality assumes correlation as well as an additional relation, i.e., the necessary condition for X. In the same way, other logical relations could be "scaled" on such a measure of complexity.

Given this kind of ideal, logical scale, different verbal constructions could be assessed in terms of their logical complexity. Concomitantly, using the same kind of scale individuals could be ordered in terms of the degree of logical complexity of their construction in different forms of behaviors (e.g., production vs. comprehension; differences between different domains of performance, such as language vs.

action). Given such independent assessments of an individual's particular forms of verbalization by this scale as well as an ordering of language utterances by this scale, one could predict comprehension of, or "adequate" performance in a given instance by a given individual, as well as predict differences in comprehension or performance based on the kind of "mismatch" of forms available for an individual for particular kinds of instances.

While we posit an increase in complexity with the inclusion of one further discrimination requirement, and each addition is assumed to increase complexity, additivity (interval scale) is not assumed in our derivation of an ordinal scale. That is, while the same differences in scale "values" may be obtained at different points in the scale, we do not hold that the degree of change need be the same at the two points--they are only different in an ordinal sense (i.e., greater or lesser than).

Again, although we have outlined the idea of a scale as if it were possible to derive such a scale completely independently of any empirical evidence, and as if we were concerned only with testing the empirical tenability of such a scale once derived, we, in fact, see the idea of positing an "ideal scale" as an heuristic device. First, the articulation of particular aspects of language behavior which may be of importance in a scale of complexity will be accomplished in the context of testing specific hypotheses derived from our previous work, and the work of others. Second, as mentioned above, instances of non-correspondence may be found between an expected order of difficulty and an observed order of difficulty in the linguistic performance of a particular group under particular circumstances. Within our general methodology for the use of this kind of heuristic scale, when such non-corresponding instances occur, requires that we look for other co-occurring variables

which could account for the discrepancy of the ideal predictions from the empirical observations.

In contrast to many other approaches using complexity (e.g., using Chomsky's transformational grammar in deriving an ordering of difficulty) to explore language comprehension behavior, the model proposed here requires that any complexity measure include at the very least:

(1) syntax (i.e., principles for combining components and associated variations in sentence form); (2) associativity among components similar to Rosenberg's concept of associativity and Perfetti's "lexical density", but not restricted to the sentence as the unit of analysis but rather the larger communication matrix; (3) number of discrete components, or length; (4) degree of correspondence between sequencing of components in "real time" and the temporal-spatial ordering of the represented event; (5) implied requirements of relational terms (e.g., temporal-spatial sequence vs. causality; coordination vs. conditionality; (6) explicitness of the communication features available to indicate variations in all of the above.

We begin with the further general assumption that (with the exception of perceptual impairments for a particular individual) the ease or difficulty of language comprehension is, in principle, equivalent in the auditory and visual modalities. Ostensible "superiority" of comprehension of auditory forms (compared with comprehension of visual forms), we hold, is usually attributable to lesser complexity of the material to be comprehended, less demanding criteria of "comprehension," or availability of some specific auditory cues not available in the visual form. Thus, if stimulus complexity and comprehension criteria are equivalent in the auditory and visual modes and the auditory form is still easier to

comprehend than the corresponding visual form, we would expect that there is available in the auditory form an explicit cue for the organization of the material which is not available in the ostensibly equivalent graphic form. That is, spoken English includes an intonational contour consisting of "melodic features" (e.g., patterns of stress, pitch, hesitation) which are neither fully explicit nor articulare in the graphic form of English.

Testing of this model in reading and listening comprehension tasks have been started. In a recent study (Oakan, 1973) attempted to achieve a "match" between characteristics of the comprehender (assessed independently) and a reading task, both assessed for the use of and the occurrence of one set of linguistic forms. Groups of "good" and "poor" readers were selected (using standard test scores and teacher ratings), carefully matching for age intelligence test scores, socioeconomic background, and classroom (6th grade). Every attempt was made to exclude any child who showed a manifest perceptual difficulty (e.g., vision; hearing) and any child who might be called English-language "deficient" (e.g., English as a second language). Each child was assessed for his language production in a story he was asked to construct for a Thematic Apperception Test card. As predicted, when the language usage of the two groups was compared, it was found that those children identified as poor readers used coordination in their organization of their sentence forms for the most part, while good readers also used sentences which included syntactical and logical subordination of components. Two sets of reading materials were then constructed, each including the same words and concepts; one set of materials was written using only coordinated components in the sentence constructions (e.g., "and;" "then"), while

the second set included subordination of sequential relations and seriation. The same comprehension questions were asked for the two forms of materials. The stories were presented both graphically and auditorily (with only intonational variations). The findings showed that the poor readers comprehended the coordinated sentence stories as well as the good readers, but showed significantly less comprehension of the subordinated materials. Further, it was found that the same pattern of results was evident for the auditory presentations as for the visual presentations for the language forms used. Whatever "difficulties" the so-called poor readers have is equally evident with auditory language when the materials are systematically controlled and varied, compared with the subjects' own spontaneous usage.

We have developed (Rubano, 1976; Wiener & Shilkret, appended here) a set of principles and a system for scaling the complexity of verbal utterances in terms of relations specified for the predictions (using connectives for the most part); that is, the scale attempts to assess the number of simultaneous relations explicitly indexed or directly inferable between predications. This scale was tested on the verbal productions of children in 4th, 6th and 8th grades (from both middle and lower socioeconomic backgrounds) describing three kinds of pictorial stimuli (movies, picture series and single pictures. First, the reliability of judges in scoring was quite high. More importantly, in terms of our concerns about a reliable and valid instrument, it was found that complexity, as expected, increased over grades, and lower socioeconomic children (presumed to be less likely to use standard English patterns) produced fewer of these explicit relations in a thought unit.

This evidence of a reliable scale, one which apparently discriminates likely language comprehension differences, makes it possible to test our match-mismatch task more directly. That is, we can assess the productions of children, and then require them to "comprehend" texts (with the same contents) when the language is of lower, of the same, and of greater complexity than that used by that child. These texts can be tested both in auditory as well as graphic forms.

We recognize that when we modify the complexity of a language instance by variations in connectives, we may also be introducing complexity in terms of syntactic organization. That is, as we have noted elsewhere (e.g., Wiener, Cormier & Shilkret, appended here) a change from "He did A and then he did B" to "Before he did B, he did A" not only adds the complexity in the connective, but presents the reader with a sequence which has a linguistic form which is different from the sequence of events in "real" time. Similarly, the use of some connectives may require sentence organization which includes additional subordination constructions or changes the "lexical density," none of which may be scored when we use this particular scale. However, despite these constraints, it would seem possible to test directly the hypotheses that (1) children comprehend a passage only to the extent that it matches the degree of complexity they already produce, and that (2) this effect is evident whether the material to be comprehended is given auditorily or graphically -- if comprehension is measured by the same criteria.

Method

Subjects

Although 72 subjects were needed for the present study, 86 6th grade students were originally administered the reading material.

These students comprised the entire 6th grade population at a public elementary school excepting those students who were non-standard English users or who had been identified as having perceptual/motor handicaps. There were 52 girls and 34 boys in this original sample.

Materials

A) Pretest: To generate a written language sample for each S, Card 1 of the Thematic Apperception Test was shown to groups of students. The card was projected onto a viewing screen in each student's home classroom.

B) Experimental procedure: Stories were constructed about three different topics (Homing Pigeons, Tornadoes and Glassmaking). Each topic, in turn, was written in forms which contained three different levels of complexity (high, middle and low), the differences being contributed most often by selected connectives. This procedure resulted in nine distinct stories.

Figure 1 insterted about here

Attempts were made in pilot research to ensure that all stories were equivalent except for differences in topic and the complexity contributed by connectives. That is, each of the three stories at the low level of complexity, for example, were written such that they were of nearly equal length, had the same number of thought units, had the same level of total complexity and that each sentence within this story level varied only minimally in connective complexity. These same controls were also applied to stories in the middle and high levels of complexity.

Figure 1. Texts of test stories at each of the three levels of complexity.

Tornadoes

Low Complexity

Tornado season is a dangerous time and these storms strike with little warning and they usually come during hot and humid weather. The direction of these black, funnel shaped clouds is tracked by radar; weather charts and wind direction are not helpful in forecasting tornadoes. Tornadoes rise and fall in a bounding motion and parts of their path are not damaged but other places are damaged badly and we see uprooted trees, demolished buildings and ruined crops. Tornadoes move forward at about 20 or 30 miles per hour but this wind speed does not cause the damage. There is an outer wind and it circles around the tornado's center. This outer wind reaches 400 miles per hour or more and it enables the inside wind to suck up objects like a vacuum cleaner. The storm moves forward a short distance and then these lifted objects are dropped elsewhere. Many people live in danger zones and there are ways they prevent damage or injury from the storms. They listen for radio warnings and when they hear that a tornado is coming, they open the windows and doors of their buildings and they put any loose objects inside and then they go to the cellar.

Number of words -- 203

Total Complexity -- 15

Number of Thought Units -- 25

Figure 1. (Continued)

Tornadoes

Middle Complexity

Tornado season is a dangerous time because these storm strike with little warning and they usually come during the hot and humid weather. Because weather charts and wind direction are not helpful in forecasting tornadoes, the direction that these black, funnel shaped cluds take is tracked by radar. Tornadoes sometimes rise and fall in a bounding motion, so parts of their paths are not damaged, but other places, where we see uprooted trees, demolished buildings and ruined crops, are badly damaged. Because the forward movement of tornadoes is about 20 or 30 miles per hour, this wind speed does not cause the damage. When the outer wind circles around the tornado's center and it reaches 400 miles per hour or more, the inside wind can then suck up objects like a vacuum cleaner. When the storm moves foward even a short distance, the objects which have been lifted may be dropped elsewhere. Many people who live in danger zones listen for radio warnings and when they hear that a tornado is coming, they can prevent damage or injury from the storms by opening the windows and the doors of buildings, putting any loose objects inside and going to the cellar.

Number of words -- 200

Total complexity -- 41

Number of Thought Units -- 25

Figure 1. (Continued)

Tornadoes

High Complexity

When tornado season comes, which is usually during times when the weather becomes hot and humid, it can be dangerous because these storms can strike without much warning. Because weather charts and wind direction would not be helpful in forecasting tornadoes, radar, which can track the direction that these black, funnel shaped clouds take, is used. Since tornadoes sometimes rise and fall with a sort of bounding motion, parts of their paths can be undamaged, and other places, where we see uprooted trees, demolished buildings and ruined crops, are damaged badly. Since the forward movement of tornadoes is about 20 or 30 miles per hour, this wind speed does not cause the damage. If the outer wind circles around the tornado's center until reaching 400 miles per hour or more, the inside wind can suck up objects like a vacuum cleaner. When the storm moves forward even a short distance, the objects which have been lifted may be dropped elsewhere. People who live in danger zones can listen for radio warnings, and if a tornado is coming, can try to prevent damage or injury by opening the windows and the doors of buildings and putting any loose objects inside before going to the cellar.

Number of words -- 204

Total Complexity -- 64

Number of Thought Units -- 25

Figure 1. (continued)

Questions - Tornadoes

- 1) The forward motion of tornadoes
 - ☐ is about 25 miles per hour and causes trees to fall.
 - ☐ does not cause any damaging winds.
 - ☐ circles around the tornado's center after picking up speed.
 - ☐ makes forecasting them difficult.

- 2) Radar tracks the direction of tornadoes; forecasting them
 - ☐ is not aided by other methods very much.
 - ☐ is never accurate.
 - ☐ requires using weather charts and wind directions.
 - ☐ depends on how quickly they spring up.

- 3) People listen for radio warnings and when a tornado is expected
 - ☐ they close the windows and doors of buildings.
 - ☐ they listen for the radar and weather reports.
 - ☐ they move to a location below ground level.
 - ☐ they fasten down all loose objects that are outside.

- 4) Objects which end up scattered in other places have been lifted up by the tornado because
 - ☐ the tornado has a bounding motion.
 - ☐ the objects were not fastened down.
 - ☐ the tornado's high-speed, inside wind sucked them up.
 - ☐ of the tornado's pointed, funnel shape.

- 5) Tornadoes spring up with little warning and usually strike
 - ☐ before any radar warnings have been given.
 - ☐ at a speed of up to 400 miles per hour.
 - ☐ only when the clouds become funnel shaped.
 - ☐ during warm, muggy conditions.

- 6) Tornadoes rise and fall in a bounding motion and this motion
 - ☐ is the cause of their great destructiveness.
 - ☐ leaves parts of their paths undamaged.
 - ☐ makes them difficult to forecast.
 - ☐ is determined by wind direction.

Figure 1. (Continued)

Homing Pigeons

Low Complexity

Homing pigeons can fly great distances and return to their home location; pigeons are started on their training when they are 28 days old. First they fly around their home nest and then they are taken to locations short distances from it and they learn to fly back to their nest from there. Their training continues and they are moved to new spots farther away. The exact reason homing pigeons return to their nest is not known, but some people think they use the sun and the stars for direction, and other people say that the birds use rivers or the length of daylight as clues. Many pigeons fly 1,000 miles to their home nest and the longest distance recorded is 2,300 miles; this distance was flown by a pigeon used by the U.S. army in carrying out a mission. Years ago, homing pigeons carried messages and they are called carrier pigeons for this reason. They were also special messengers in times of emergency and were important for this task; when an army was in battle, carrier pigeons could fly out unnoticed. Today, some people keep these birds for hobbies and they join homing pigeon clubs and enter their birds in races.

Number of words -- 203

Total complexity -- 17

Number of Thought Units -- 24

Figure 1. (Continued)

Homing Pigeons

Middle Complexity

Homing pigeons can fly great distances and return to their home location because they are started on their training when they are 28 days old. After they fly around their home next, they are taken to locations short distances from it and they learn to fly back to their next from there, and as their training continues, they are moved to new spots farther away. The exact reason homing pigeons are able to return to their nest is not known, but some people think they use the sun and the stars for direction, while other people say that the birds use rivers or the length of daylight as clues. Many pigeons are able to fly 1,000 miles to their home nest and the longest distance ever recorded is 2,300 miles. This distance was flown by a pigeon used by the U.S. army to carry out a mission. Because homing pigeons used to carry messages years ago, they are sometimes called carrier pigeons. They were also special messengers during times of emergency, as when an army was in battle, and were important for this task. They could often fly out unnoticed. Today, because some people keep these birds for hobbies, they join homing pigeon clubs and they enter them in races.

Number of words -- 209

Total Complexity -- 42

Number of Thought Units -- 24

Figure 1. (Continued)

Homing Pigeons

High Complexity

Because they are started in training when they are 28 days old, homing pigeons can return to their home location after flying great distances. Not until they fly back to their home nest after being taken short distances away, are they moved to new spots farther away. Otherwise, they can only fly around their home nest to continue training. While the exact reason homing pigeons are able to return to their nest is unknown, some people think they do this by using the sun and stars to tell directions; other people say the birds use rivers or the length of daylight as clues. Many pigeons are able to fly as far as 1,000 miles back home, and the longest distance ever recorded is 2,300 miles which was flown by a pigeon used by the U.S. army to carry out a mission. Because homing pigeons were frequently used to carry messages years ago, they are sometimes called carrier pigeons. They were important, because if an army were in battle, they could be used like special messengers in such times of emergency since they could often fly out unnoticed. Today, because some people still keep these birds for hobbies, they often join clubs so they can enter them in races.

Number of words -- 207

Total Complexity -- 68

Number of Thought Units -- 24

Figure 1. (continued)

Questions - Homing Pigeons

- 1) Many pigeons fly 1,000 miles. The record is 2,300 miles and that pigeon
 - ☐ was used to carry messages for an army in battle.
 - ☐ helped people in an emergency.
 - ☐ helped the U.S. army in a mission.
 - ☐ recorded this distance in a competition.

- 2) While in training, young homing pigeons are kept close to their home nest
 - ☐ so that they can fly out for long distances.
 - ☐ and are trained to carry messages.
 - ☐ for the first 28 days of their training.
 - ☐ and first learn to fly back to it from spots close by.

- 3) Years ago, homing pigeons were important to people because
 - ☐ they could fly out undetected in emergencies.
 - ☐ they helped people communicate in emergencies.
 - ☐ they were the only means of communication.
 - ☐ they had the speed for carrying messages in battle.

- 4) At the present time, people sometimes keep homing pigeons and they
 - ☐ train them from birth as a hobby.
 - ☐ have them take part in races.
 - ☐ teach them to carry messages.
 - ☐ teach them to use rivers and the length of daylight to return to their home nest.

- 5) Homing pigeons fly great distances and return to their home location. They
 - ☐ begin learning this when they are about one month old.
 - ☐ learn this first by flying from distant spots.
 - ☐ do this as a hobby.
 - ☐ usually carry messages.

- 6) Some people think that homing pigeons use rivers or the amount of daylight to find their way back home, but
 - ☐ they can usually do this only for distances of less than 1,000 miles.
 - ☐ they sometimes fly back to their distant locations instead.
 - ☐ they actually use a combination of stars, sun, rivers and daylight.
 - ☐ nobody really knows how they do it.

Figure 1. (Continued)

Glass

Low Complexity

Glass-makers from long ago would be surprised to see the many uses of glass today. In America, they first made glass in a Virginia factory over 350 years ago and glass was used mainly for beauty; they made two kinds of glass, one kind for beads and one kind for bottles. Now, more than 50,000 different kinds of glass are made and different "recipes" are used for each one. Glass-makers are sawing, floating, hammering and bending glass and they create brand new types of glass and can help supply many needs of different buyers. Certain recipes make glass stronger than steel, other glass is spun into threads and it looks like cotton. Glass is made mostly from sand and colored glass is manufactured by adding certain metals when the glass is liquid. Glassblowers make special shapes and figures and they create them by blowing puffs of air through long tubes and many people think this kind of glassblowing is artistic. Most glass today is inexpensive; it is made in factories in big furnaces and with pressing machines, and it is possible to use old, discarded glass and make new glass from it and this process requires less energy from the melting furnace.

Number of words -- 204

Total Complexity -- 15

Number of Thought Units -- 24

Figure 1. (Continued)

Glass

Middle Complexity

Early glass-makers, who first began making glass in America in a Virginia factory over 350 years ago, would be surprised to see the many uses of glass today. Because glass was used mainly for beauty, two kinds of glass were made, for beads and for bottles. More than 50,000 kinds of glass are now made by using different "recipes." Glass-makers are sawing, floating, hammering and bending glass to help supply the many needs of different buyers and they also create brand new types of glass. Glass that is stronger than steel is made by using certain recipes, other glass is spun into threads and it almost looks like cotton. Glass is made mostly from sand and colored glass is manufactured by adding certain kinds of metals while the glass is still liquid. If glass-blowers blow puffs of air through long tubes, they create special shapes and figures which is thought to be artistic by many people. Because it is made in big furnaces and with pressing machines in factories, it is possible to use old, discarded glass to make new glass from it. Most glass today is inexpensive because it requires less energy from the melting furnace.

Number of words -- 200

Total Complexity -- 41

Number of Thought Units -- 24

Figure 1. (Continued)

Glass

High Complexity

If early glass-makers, who first began making glass in America in a Virginia factory over 350 years ago, could see the many uses glass has today, they would be surprised. Because glass was used mainly for beauty, two kinds of glass were made, for beads and for bottles. More than 50,000 kinds of glass can now be made by using different "recipes" and glass-makers can saw, float, hammer or bend glass to help supply the many needs which different buyers have while creating brand new types of glass. Glass that is stronger than steel can be made by using certain recipes whereas other glass is spun into threads to make it look almost like cotton. Colored glass can be manufactured if certain kinds of metals are added while the glass, which is made mostly from sand, is still liquid. If glass-blowers blow puffs of air through long tubes, they can still create special shapes and figures which is thought to be artistic by many. Because it is made in big furnaces and with pressing machines which are in factories, it has become possible to use glass that is old and discarded to make new glass. Most glass is inexpensive since it requires less energy from the furnace.

Number of words -- 207

Total Complexity -- 69

Number of Thought Units -- 24

Figure 1. (continued)

Questions - Glass

- 1) Colored glass can be made when certain metals
 - ☐ are added instead of sand.
 - ☐ are melted before going into the furnace.
 - ☐ are melted along with recycled glass.
 - ☐ are added when the glass is melted.

- 2) Today, glass is produced in many different ways
 - ☐ but can only be done by artistic glass-blowers.
 - ☐ and gives buyers the special kinds they want.
 - ☐ requires puffs of air for their shape.
 - ☐ and is made just like they used to do 350 years ago.

- 3) Old discarded glass can sometimes be made into new glass and this process
 - ☐ uses less power than making new glass from sand.
 - ☐ is performed in big factories.
 - ☐ requires just one special recipe.
 - ☐ uses pressing machines.

- 4) Today, the creation of special shapes and figures by glass-blowers is
 - ☐ appreciated as an art.
 - ☐ a very expensive process.
 - ☐ done the same way as years ago.
 - ☐ becoming less frequent as recycling takes over.

- 5) Early glass-makers would be most amazed at
 - ☐ how much less time is needed to make glass.
 - ☐ how recipes are used for glass-making.
 - ☐ glass being stronger than steel.
 - ☐ the number of things glass is used for.

- 6) Some glass looks almost like cotton. This is done
 - ☐ by injecting tiny air bubbles into the glass.
 - ☐ by using special furnaces.
 - ☐ by spinning it into threads.
 - ☐ by glass-blowers.

Table I
Ranges of Length, Thought Units and Complexity
for the Three Levels of Complexity

Complexity Level	Total Words	Total Thought Units	Total Complexity Score
Low	203-204	24-25	15-17
Middle	200-209	24-25	41-42
High	204-207	24-25	64-69

All nine stories were also tape recorded on cassette tapes so that these stimulus paragraphs could be presented in an auditory mode to half of our subjects. A tutor in the school system with several years of elementary classroom experience read the stories for the tape recording. Intonational variations and stress patterns which were shown to aid comprehension (Shilkret & Wiener, 1974) were carefully minimized (except for an end-of-sentence fall in pitch, so that the graphic and auditory modes would be more comparable in terms of the number of comprehension aids in the material itself).

Booklets were constructed for the reading task so that the presentation order of story topic was held constant (i.e., all Ss read stories about tornadoes, homing pigeons and glassmaking in a fixed order). However, the presentation of complexity was counterbalanced. Six different orders were necessary, resulting in six or seven Ss being tested on each order. In the booklet, each story was followed by two pages of blank, lined paper and a page of the six multiple choice questions corresponding to that topic.

Again, pilot work showed that the multiple choice questions used in this study could not be answered correctly at better than chance level by Ss who had not read the stories. Questions had also been reviewed to be sure that no single question contributed disproportionately to either correct or incorrect response totals by pilot subjects.

Procedure and Instructions

A general attempt was made throughout the study to produce an attitude of interest and cooperation in the subjects.

All responses to the TAT card were gathered in one day. Moving from one class to the next, the experimenter was introduced by the classroom teacher and allowed to use approximately 15 minutes of class-time to gather this data.

The following introduction was given to all students:

I am very interested in the way kids your age write stories and how you read stories. I'd like to find out what things you write about and what you think about stories that I have made up. Other students have helped me with this and they all seemed to enjoy it. I hope you do too. Remember, what we do today has nothing to do with your school grades or marks. This isn't a test but a way for me to find out more about the way kids read and write.

Next, after distributing some lined paper and arranging the projector and screen, the following instructions were read to the class:

First, we'll start with the story-writing part. Here's how we'll do it. I have a picture to show everyone. Your job will be to look at the picture and then make up a story about it. This picture has a person in it so you can write about what the person is doing, what happened before this and what happened after. Write your story so it has a beginning, a middle and an end. You could also write about what the person or people in your story are thinking and feeling. Otherwise your story can be about absolutely anything you want. Whatever comes to your mind can be put in the story. They can be funny or sad, serious or not serious, it's up to you. Any questions? OK, here we go.

After about two minutes, Ss were reminded that their stories should tell about previous and future events and how they relate to the picture. Ss were given 5 to 8 minutes to complete their stories. The class was then reminded that they would complete this study the following week when the experimenter would again visit their class.

Four separate school visits were required to complete administration of the reading tests. These were made within the course of nine days. Completion of the reading tests took an average of 45 minutes. An attempt was made to have approximately half of each classroom take the graphic mode and the other half the auditory mode of the reading tests. Since the graphic material was ordered and arranged in booklet form beforehand, larger numbers of Ss could be administered the tests in the graphic mode, while counterbalancing requirements limited auditory presentations to groups of seven. The school library and the tutoring room were used for the administration of both modes of the reading test.

Graphic mode

Students were seated comfortably at desks or tables and spaced at least one seat width apart. The following instructions were read to them:

Now I'm going to show you some stories that I made up. I'm trying to test out how easy these stories are to understand. There are three of them to read. After each one, I'd like you to first write down as much as you can remember from the story. You'll use the lined paper in the booklets. We will wait for everyone to finish this part; then, there are some questions to answer so that I can get an idea of how well you understood it. They will be regular multiple choice questions like these (show sample sheet of questions), so all you have to do is check off the best answer. I will read each question and the four possible answers out loud while you can follow on the page. Be sure to read the printed stories carefully, because when you are through, you won't be able to turn back to the story. So there are two things you will do after reading each story: first write down as much as you can remember about it, and then wait until everyone is finished; second, answer the multiple choice questions after I've read them with you. Everybody ready?

Five minutes was allowed for the S who finished reading the material last to complete the recall task. No subject was allowed to turn to the multiple choice questions before this time. All Ss turned to the questions together; each question and the four multiple choices for each were read aloud by E.

Auditory mode

Ss who were administered the reading tasks in the auditory mode were also given a booklet of blank, lined paper and the multiple choice questions. Ss were seated around a portable cassette recorder. The six counterbalanced orders of presentation were arranged in advance and played to six separate groups of seven Ss. Instructions for the auditory mode closely paralleled those for the graphic mode. Obvious changes such as reference to "hearing" the story rather than "reading" it were made, but instructions were otherwise identical for both conditions. Similarly, conditions for the recall and response to the multiple choice questions were exactly the same as for the graphic mode.

Results

Pretest Scoring

Three judges, independently of each other, scored the 82 stories produced by the children for the TAT card. The criteria for scoring were those specified and elaborated by Rubano (1976) (see Rubano, Wiener & Shilkret in this report). For this scoring, two judges agreed on 80.4% of all trails with a third judge differing from 1 to three scale units. Overall the scoring of 96% of the protocols showed a difference of 1 or less for any two judges.

The three judges then went over all items with any disagreement among the judges; these instances were discussed to resolve any differences which could be considered a function of non-agreement about the

length of thought unit and/or the scoring values of a particular construction. Once these scoring problems were resolved, a subject received the highest score agreed upon by at least 2 of the judges.

The subjects were assigned to the (1) Low Complexity group, scores 0-4; the (2) Middle Complexity group, scores 5-7; or (3) the High Complexity group, scores of 8 and above. The comprehension data were analyzed for both (1) number of questions answered for each story, and (2) the number of correct recalls for each story. These sets of scores were analyzed in an Analysis of Variance design with three levels of subjects (Low, Medium and High); two types of tasks (Auditory and Visual); and three levels of stories (Low, Medium and High) as a repeated measurement. Table I shows the summary of the analysis for the number of errors on the comprehension task. Only the data for 79 subjects could be analyzed; the remaining three did not complete the comprehension tasks.

Table II

Summary of the Analysis of Variance, Using Three Levels of Subjects, Two Modes of Presentation and Three Degrees of Complexity of Text, for the Number of Errors in Responses to the Comprehension Questions

Source	df	M.S.	F	p
<u>Between subjects</u>				
Level of Complexity	2	2.72	1.02	NS
Auditory vs. Visual	1	14.20	5.34	.01
Level of Complexity by Mode	2	10.68	4.01	.05
<u>Within subjects</u>				
Story level	2	.56	<1	NS
Level of subject x Level of story	4	1.40	<1	NS
Mode by level of story	2	.54	<1	NS
Mode x Level of story x Level of subject	4	1.12	<1	NS
Subject within subject group	146	1.42		

Examination of the means show that in contrast to many investigator's expectations, the auditory mode of presentation showed a higher means errors score than for the visual mode ($F_{1,73} = 5.34, p < .01, M_V = 2.60, M_A = 3.21$). The significant interaction of the level of complexity of a subject's production and the mode of presentation of text ($F_{2,146} = 4.01, p < .05$) shows the following distribution of error scores:

	<u>Auditory</u>	<u>Visual</u>
$M_{\text{level 1 subjects}}$	2.94	3.27
$M_{\text{level 2 subjects}}$	3.07	2.37
$M_{\text{level 3 subjects}}$	3.54	2.39
Total	3.21	2.60

Exploration of these means show that subjects who used the least complex verbal form in their production did better on the auditory mode and less well on the visual mode, while the other two groups show the opposite results. In addition, the highest level of complexity producers do least well of all subjects on the auditory mode.

A similar analysis of the number of thought units recalled (either (1) exact paraphrase, (2) exact paraphrase with minor non-essential omission, or (3) less explicit recall but clearly appropriate) showed non-significant results for the several factors analyzed, except that subject groups (low, medium & high) differed ($F_{2,73} = 2.63, p < .08$), with the middle group recalling more than either of the other groups, independent of mode (although the pattern of means for auditory mode showing less recall by higher level subjects than for the visual mode, but the opposite for lower level subjects).

The mean correct recall scores are as follows:

	<u>Auditory</u>	<u>Visual</u>	<u>Total (A+V)</u>
$M_{\text{level 1 subjects}}$	5.78	5.17	5.54
$M_{\text{level 2 subjects}}$	6.11	7.70	7.17
$M_{\text{level 3 subjects}}$	5.81	6.18	5.96

Nothing in the analyses of these data show any significant relationship between the level of subject production of complexity and the comprehension of texts of different levels of complexity, measured by the comprehension questions or the recall scores. What does show is that all subjects produced more complex utterances ($F_{2,146}=5.73$, $p<.01$) for the middle and high complexity level text (independently of auditory or visual modality) relative to the lowest level of text complexity ($M_L=3.29$; $M_M=4.37$ and $M_H=4.25$).

Discussion

While there appears to be some interesting findings, e.g., that for subjects who produce much verbal complexity the visual mode is associated with more correct comprehension scores than is the auditory mode of presentation, the opposite seems to be the case for the subjects who produce low levels of verbal complexity. The failure to find any data in support of our general hypotheses of a predicted correspondence between comprehension and verbal productivity scores is highly surprising. Oakan, who had used variation in coordinates vs. subordinate clause constructions and had found a comprehension difference for good and poor readers, had constructed his texts in a similar although less systematic way. Several

explanations for the differences in his results from ours may seem reasonable; each of these explanations which could account for the differences in findings requires some further systematic exploration.

First, the texts were unfamiliar contents for all of our subjects in this study, relative to Oakan's tests. Cormier, Wiener & Shilkret had noted that "unfamiliar" contents are more difficult for subjects than texts about familiar events in the world. However, the percentage of items recalled for the texts in this study was not significantly lower than those recalled in the Cormier, Wiener & Shilkret study for familiar content; in fact, if anything, our subjects in this study recalled a slightly higher percentage of the possible number to be recalled.

The more critical issue for us, was the low number of correct test scores on the comprehension test over all texts. All subjects for all conditions, including the very easiest texts showed a correct score of no more than 3 correct out of a possible 6, for any story. This low number of correct responses suggests that our questions may have been so difficult to obscure any possible text effects. It is hoped that we can develop comprehension questions which will show at least 80% correct trials for an easy text before we can hope to show any effects of more difficult texts; that is, if complexity of verbalization in a text is expected to increase the difficulty of a text. Certainly the literature is replete with data that indicate more complex grammatical constructions are expected to increase text difficulty and it is clear that our text variations seem to have introduced some grammatical text complexity. However, as noted here, any difficulty introduced by such changes may have been obscured by other sources of difficulty, in this case, the comprehension questions.

Incidentally, again, given that the recall score noted only items (not their relationships) any comprehension at all of the texts would be expected to have produced the same amounts of recall. We plan on studying further those items recalled by our subjects versus those not recalled, to determine which are "easy" and which are "difficult" in terms of text difficulty (if there are any differences in which ones are recalled as a function of text difficulty).

Clearly more research seems indicated, if we are to maintain that verbal productivity predicts comprehension, at least in terms of the scale we have applied to assess our readers and our texts.

Effects of redundancy and intonational stress in facilitating
children's comprehension of temporally-ordered events¹

Robert B. Shilkret, Marian Viets, and Morton Wiener

Mount Holyoke College

Clark University

Occasionally, in a conversation between an adult and a child, when the child fails to comprehend something, the adult will repeat what he or she originally said, but with some interesting modifications that often result in the child's now comprehending the utterance. For example, the adult may emphasize particular words of importance, introduce new words redundant with information in the original utterance, repeat a word, speak more slowly, in less complex or shorter sentences, and so on. These changes in the stimulus are important, since they seem to suggest ways in which a child may comprehend, and eventually produce, a new form. Our research (Wiener & Shilkret, 1974) has been investigating under more controlled experimental conditions some of the variables that may facilitate performance. The general logic of our method rests on the assumption that if conditions can be arranged so it will be possible for a child at one level of language usage to function as if he or she were at a more advanced level, it can then be inferred what must occur during the natural transition from one level to another during acquisition.

In the present study, we chose to investigate the young child's comprehension and production of sentences expressing temporally-ordered events. To express two events, one could say, for example, "Do A first, then do B second", or, in a more complex form, "Do B after A". Words like "before" and "after" make possible an "inverted" sentence, in which the order of lin-

guistic elements does not correspond with the order of events in "real time". Such inverted sentences, while extremely efficient in expression, have typically been found not to be comprehended by young children. In this study, it was assumed that sentences with the linguistic order of elements corresponding with the real-time order of events (that is, "non-inverted" sentences, such as "After A, B"; "A before B"; "A and then B") are less "complex" than corresponding sentences in which this linguistic order is inverted (for example, "B after A"; "Before B, A"; "B, but first A").

It was expected that increasing redundancy--that is, the number of times the temporal relation was mentioned in the sentence--would facilitate comprehension of such inverted sentences (for example, a sentence like "B after A" was expected to be harder than one like "B after A first").

Further, we expected that intonationally emphasizing or stressing the important temporal words would facilitate comprehension, particularly for the more difficult sentences. Some of our earlier work (Shilkret & Wiener, 1972) has suggested that intonational contour facilitates comprehension efficiency, particularly for syntactically complex sentences, as compared with simpler forms; and for some children who seem to have difficulty imposing organization on language in the written form (i.e., certain children labeled "poor readers"), intonational contour seems especially important in auditory comprehension. Another recent study (Shilkret, 1974), investigating adults' comprehension speed of active and passive sentences, found that emphatic stress

placed on the logical actor of passive sentences makes such sentences apparently as easy to comprehend as their active sentence counterparts.

Subjects for this study were 32 English-speaking middle class children with no known speech or hearing difficulties. Average age of the subjects was four and one-half years, with a range from 3 years, 5 months to 5 years, 2 months.

The experiment consisted of two kinds of tasks, which all children performed, a comprehension task and a structured production task.

In the comprehension task, the child was asked to "act out" each of twenty pre-recorded sentences with a set of toys. The sentences for this task all described two events performed by a single actor; events were chosen which were equally plausible in either of two possible orders of occurrence. That is, a sentence like "He brushed his teeth before he went to bed" was not used, since such a sentence describes two events which are more likely to occur in that order than in the reverse order. Studies by us and others have suggested that "event plausibility" is an important variable in comprehension of sentences of various syntactic types, particularly for young children.

The entire set of sentences for the comprehension task was constructed in accordance with three variables. First, half the sentences had emphatic stress on the temporal word or words (e.g., "After the girl patted the zebra, she rode the elephant"), while the other half were recorded with normal intonational contour. Secondly, in half the sentences, the linguistic sequence corre-

sponded with the order of occurrence of events (e.g., sentences like "After the cow ate some grass, she drank from the lake" and "The giraffe ate from the tree before he climbed over the box"), while in the other half of the sentences, the linguistic sequence inverted the order of occurrence of events (e.g., sentences like "Before the giraffe ate from the tree, he climbed over the box" and "The elephant climbed over the box after he pushed the log"). Thirdly, the sentences were constructed to vary in the amount of redundancy of temporal information in the sentence. The word redundancy here is used to refer simply to the number of temporal words in the sentence, and is not used in the strict sense of that term. Some sentences, like the ones mentioned earlier, had only one temporal word; other sentences of this first level of redundancy used the "and then" and the "but first" constructions. Sentences at a second level of redundancy included two temporal words (e.g., "Before the horse ate some grass he first jumped the fence"--here, an inverted sentence). Other sentences had three temporal words (e.g., "The pig rolled in the mud first and then he climbed over the fence next"--this is a non-inverted sentence). Each of these three variables and the combinations of them were represented orthogonally among the set of sentences. Subjects and predicates (e.g., zebra, duck; drank from the lake, walked to the tree) were intermixed across stimulus conditions. Each subject worked with half the sentences thus generated, presented in a fixed random order; the three stimulus variables were counterbalanced across trials and across subjects.

The second part of the experiment was designed to sample

the child's verbal production of these kinds of sentences. This was done by, in effect, "reversing roles" of the child and the experimenter--here, the experimenter acted out, with a different set of toys, two-event sequences, asking the child each time to describe what happened. This production task consisted of three parts, each done twice with a different set of toys. First, the experimenter acted out a "two-event" sequence (such as enacting the sequence of a zebra running across a field and then jumping over a box) and simply asked the child to describe what had happened. Next, the child was shown another sequence of events, acted out with toys by the experimenter, and asked to complete a sentence, after the experimenter provided a sentence stem. The stem was the second event of the series, with no temporal word (in the example of a zebra running across a field and jumping over a box, the experimenter said, "The zebra jumped over the box _____"). For the third part of the task, as before, the experimenter enacted the events, and provided a sentence stem to be completed by the child, but here the sentence stem, again referring to the second event, had an appropriate temporal word (in the same example, the stem was "The zebra jumped over the box after ____" or "Before the zebra jumped over the box ____"). This was done twice, once using "after" and once using "before". Actors and actions used for this part of the task had not been used in the earlier comprehension task, and the three parts of the production task were administered in the same order for all children. We thought that the first part, the free description, would be relatively easy and would elicit non-inverted descrip-

tions, that the second part would be the most difficult, and that the third part, with the sentence stem containing an appropriate temporal word, would be of intermediate difficulty, in that it required the child simply to remember the first event in completing the prompted inverted sentence.

Children's performance on both tasks was scored by two procedures. One scoring system noted whether each trial was correct or incorrect. For example, on the comprehension task, a response was regarded as incorrect if the child did not act the events in the correct sequence (that is, if he reversed the order of events, performed only one event, or included within his performance an event which was not part of the stimulus sentence). A second scoring system was developed to take into account the accuracy of response, as in the first scoring system, as well as the efficiency of response (that is, whether or not the child hesitated before responding, repeated the sentence, or asked that the sentence be repeated), and the type of error, if any. Independent sets of analyses were performed on the accuracy scores and the more finely-graded response measures; since these analyses yielded essentially the same pattern of results, only the former, results for the accuracy scores, will be reported here. Only differences of a reliability of .05 or less will be reported here.

To assess performance in the comprehension task, a three-factor mixed design analysis of variance, with repeated measures, was performed on the mean accuracy scores, expressed as percentages correct. It was found that sentences with stress on the temporal words were easier (87% correct) than sentences with no

such stress (74% correct). We had expected that emphatic stress would be particularly helpful in facilitating comprehension of the hardest sentences, that is, the inverted sentences with only one temporal word. No such three-way interaction was found; rather, stress appeared to facilitate comprehension of all stimulus conditions equally.

The analysis also revealed a significant interaction between the inversion-noninversion variable and the redundancy variable. The six means composing this interaction were compared systematically, using the t-test for differences among several means. It was found that inverted sentences with only one temporal word were significantly more difficult than each of the other five conditions. Looking separately at the non-inverted sentences, that is, sentences in which the order of phrases corresponded with the order of events, it was found that these sentences were relatively easy to comprehend by these children (approximately 84% correct), and that adding temporal words did not increase ease of comprehension. Inverted sentences with only one temporal word (e.g., "after"; "before") were harder to comprehend than their non-inverted counterparts. However, adding one other temporal word to these sentences, or, in the terminology here, making the temporal information more redundant, made the inverted sentences as easy to comprehend as non-inverted sentences. Adding a third temporal word did not substantially improve performance, as compared with second-level redundancy sentences.

This pattern of results is important, in that it suggests

that children of this age do not manifest a general inability to comprehend inverted sentences. A sentence like "The girl rode on the elephant after she first patted the zebra" was comprehended as well as any other sentence among the stimuli. A further question is how a child comprehends such a sentence, given the same child's failure to comprehend other kinds of inverted sentences. That is, does the child correctly comprehend the word "after", the word "first", or both words in such a sentence? Perhaps it is only for this third possibility that a characterization of "redundancy" of temporal information is fully appropriate. Unfortunately, it is not possible to answer this question directly from the sentences used in this study. In another recent study, Amidon and Carey (1972) found that sentences containing "first" and "last" were easily comprehended by young children, as compared with "before" and "after" inverted sentences. In that study, however, it would seem that number of temporal words was confounded across sentences. It is reasonable to expect that "first" and "last" are easier to comprehend than "before" and "after"; the former can be used in sentences with two independent clauses, whereas the latter, when the only temporal information in the sentence, require a relation between two events to be understood. In terms of the findings of the present study, this issue is not clearcut: inverted sentences with "first" as the only temporal word were still somewhat more difficult to comprehend than non-inverted sentences. A study investigating how such sentences are comprehended is currently in progress, using a method that will, hopefully, uncover the subjects' strategies of comprehension. Furthermore, results here suggest a possible training

procedure for comprehending the more difficult of the inverted sentences. Adding other temporal words and emphasizing them intonationally both seem to facilitate comprehension; removing such stimulus cues after the child has demonstrated mastery of inverted sentences may make it possible for the young child to comprehend an inverted sentence with only one temporal word.

Performance on the comprehension task was examined separately in terms of the kind of temporal word used--"before" versus "after". A two-way analysis of variance was performed on the accuracy scores, with words and inversion-noninversion as the variables. It was found that, overall, sentences with "before" were comprehended more easily than sentences with "after"; this difference was accounted for almost entirely by the relatively lower performance on the inverted sentences with "after", as compared with the other three conditions, which did not differ from each other. It might be noted that in these sentences, the "after" inverted sentences, the subordinate clause introduced by "after" appears after the main clause. An independent analysis indicated that, for all sentences in the study, sentences with a temporal word in the initial position of the sentence were easier than sentences with the temporal word in a delayed position.

Let us turn now to these subjects' performance on the structured production task. The free description items were the easier--and here, the majority of subjects (70%) produced fully correct sentences, with the order of mention congruent with the order of events. In a sense, success on this part of

the task is another kind of indication that children this age have little difficulty remembering two events and expressing them linguistically. The analysis of variance indicated that performance on the second part of the task--the one requiring completion of a sentence stem which represented the second event--was quite difficult for these children. On the other hand, the third part of the task--completion of a sentence stem representing the second event and including an appropriate temporal word--was significantly easier. In fact, under these relatively structured conditions, these children completed the stems, thus producing inverted sentences, with as great a facility as they produced noninverted sentences in the free description part of the task. A separate t-test for correlated measures on the two items of the third part of the task indicated no apparent difference in performance between the "before" stem and the "after" stem.

The kinds of errors children made in completing the sentence stems is revealing. A typical kind of error was a reversal, but occasionally the child's spontaneous comments suggested an ability not fully captured by the performance on the task alone. One subject, for example, on the second part of the task, when the experimenter enacted a toy cow eating grass and then walking around a toy lake, and given the stem, "The cow walked around the lake ____", replied, "And then he ate some grass". This completion is incorrect, since it reverses the sequence of events. Immediately after giving her response, however, the child told the experimenter that the sentence had been "backwards the first time." She explained, "Because, because you said he walked around the pond, and I said he ate some grass, but that's backwards". This

child seemed "aware" of the correct temporal relation between events--that her production was not correct--but unable to articulate the inversion when required to do so. On the third part of the task, however, when provided with an appropriate word to express the inversion, this child quickly completed the stem correctly, and seemed aware of her success here.

Another kind of error seems to represent a slightly more advanced approach to the problem of completing an inverted sentence. To the same item, seeing a cow eat some grass and then walk around a lake, and hearing "The cow walked around the lake___", one child replied hesitantly, "And then . . . He went to the grass and then walked around the lake". This is an accurate description of the sequence of events, complete with a temporal index perhaps indicating the relationship, but it is not a completion of the inversion. This child, again, seemed to understand the relationship of the events as well as the difficulty of the task, and, in a sense, solved the problem by beginning a new series of events which resulted in the actor doing one of them twice. By the way, both these children demonstrated the ability to comprehend inverted sentences in the earlier part of the experiment.

In summary, the results here seem to suggest some stimulus conditions, like intonational emphasis and redundancy of temporal information, which may facilitate the young child's comprehension of a relatively complex event, and enable the child to perform at a level beyond his normal level of language usage. It has recently been said that there are really only two theories

of language acquisition, an "impossible" theory and a "miracle" theory. Nonetheless, language acquisition does seem to occur despite the seeming impossible complexity of that fact. The focus of this study, and of our broader program of research investigating the use of logical relational terms, is more on specifying and predicting conditions under which successes and failures occur than on those successes and failures alone. It is equally as important, from our point of view, to look at the adult's modification of his or her speech to a child, and how this modification changes during language acquisition, often resulting in the child's understanding or saying something that the child had not before, as it is to look at the child's speech itself.

FOOTNOTES

1. Paper presented at the Annual Meeting of the Linguistic Society of America, San Francisco, December, 1975. For reprints, write Shilkret, Department of Psychology, Mount Holyoke College, South Hadley, Mass. 01075. The study reported as a part of this research was M. Viets' 1975 Honors Thesis at Mount Holyoke College. Some of the research discussed here was supported, in part, by grants to Wiener and Shilkret from the National Institute of Education, and to Shilkret from Mount Holyoke College.

REFERENCES

1. Amidon, Arlene, & Carey, Peter. Why five-year-olds cannot understand before and after. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 417-423.
2. Shilkret, Robert B. Intonational contour and language comprehension: Intonational stress as a variable in comprehending sentences of different forms. Unpublished Doctoral Dissertation, Clark University, Worcester, Massachusetts, 1974. (Microfilm No. 75-12,222)
3. Shilkret, R., & Wiener, M. The contribution of syntactic and parasyntactic cues in the comprehension of spoken and written language. U.S. Office of Education Research Grant, Final Report, February, 1972 (available from author).
4. Wiener, M., & Shilkret, R. Complexity in auditory and graphic language. National Institute of Education Research Grant, 1974 (Grant No. NIE-G-74-0017).

Comprehension of sentences expressing three events and the temporal
relations between them¹

Patricia Arns

and

Robert Shilkret

Mount Holyoke College

The production and comprehension of sentences representing temporal relations between events have recently been used as a paradigm of more general language performance. Words like "before" and "after" allow one to represent a series of events in ways which may or may not correspond with the "real time" order of events. For example, in the sentence "Open the door after you answer the phone", one is instructed to perform first the task mentioned last, namely, answering of the phone. The degree of such correspondence or non-correspondence between the order of clauses and order of events those clauses represent has been used as one index of complexity or predicted sentence difficulty (e.g., Clark, 1970, 1971, 1973; Clark & Clark, 1968). Other proposed general principles of language comprehension have been invoked, using these kinds of sentences to test those principles. For example, Smith and McMahon (1970) proposed that information in the main clause is more easily interpreted than information in the subordinate clause, a principle which is orthogonal to the order-of-mention hypothesis. Another analysis, that of Bever (1970) has proposed that the "assertion" of a sentence is the "basis" for understanding, with the "presupposition" being "subsidiary" to it. For two-event sentences with "before" or "after", these principles lead to the same predictions, one of them being that "before sentences are easier than "after" sent-

ences since in "before" sentences "the assertion, which is in the main clause, also describes the first event, whereas in "after" sentences the assertion describes the second event" (Amidon & Carey, 1972, p. 418). The Clarks' analysis also leads to the prediction that "before" sentences should be easier than "after" sentences, but for a different reason; by their analysis, word meanings are learned as a hierarchy of semantic components. However, as recently noted by Bloom (1975), the statement that a particular feature like [+prior] is learned before a feature like [-prior] is "more a paraphrase [a description of results] than it is an explanation" (p. 272). The Clarks also use a principle of "derivational simplicity" to describe experimental results for comprehension and production tasks with two-event sentences. This principle, which is similar to those of Smith & McMahon and Bever, maintains that sentences in which the main clause precedes the subordinate clause are easier to comprehend than those in which the subordinate clause precedes the main clause. An added difficulty is the lack of grounds for knowing which of two or more independent principles of analysis is taken to be more important in predicting sentence difficulty.

Many of the studies in this area have examined young children's comprehension and production, and the results have been mixed. For example, Amidon & Carey (1972) and more recently Friedman & Seely (1976) did not find a difference between "before" and "after" sentences in young children's comprehension; Barrie-Blackley (1973) found "after" sentences to be easier. Several

studies (e.g., Hatch, 1971; Amidon & Carey, 1972; Shilkret, Viets, & Wiener, 1975) have found that inverted sentences, in which the order of mention does not correspond with the order of events, appeared to be harder than non-inverted sentences, in which there is a correspondence between these two orders. However, this does not seem to be due to a general inability at a given developmental period to comprehend inverted sentences, as might be predicted from a "pure" semantic features analysis. Amidon & Carey and Shilkret, Viets, & Wiener found that inverted sentences with "first" and "last" as the temporal indices were comprehended as easily as non-inverted sentences; additionally, the latter study found that emphasizing the temporal words by intonational stress facilitated the comprehension of the "before" and "after" sentences for young children.

Most of the studies in this area have used two-event sentences, thus restricting the number of sentence types that can be generated. Such sentences do not allow for certain types of transformations--either the order presentation of events in a given sentence agrees with the order of occurrence of events, or it does not. Often, this does not allow for an assessment between independent principles of proposed complexity. Therefore, in the present study, three-event sentences (e.g., "The dog ate the food before he drank the water and before he entered the house") were used to allow a wider range of inversions and transformations, and thus to allow an assessment of the generality of proposed principles of complexity as well as a comparison

between independent principles (e.g., order of mention vs. derivational simplicity). Sentences which directly indicate the temporal locus of events, without the use of subordinating conjunctions (e.g., "First the girl read a book, second she answered the phone, and third she watered the plants"), were also included as a "baseline" measure of comprehension of various transpositions required to repeat the events in their "real time" sequence. These "baseline" sentences control for the effect of subordinate constructions, since, in effect, each clause in such a sentence is a main clause (Amidon & Carey, 1972).

Three response measures were used here, since the obtained ordering of difficulty between sentence types may well differ, depending on the response measure taken. Most studies have used a measure of number of correct sentences and/or number and types of errors. Smith & McMahon's (1970) study with adult subjects also included a latency measure, and that was included here as well. We also used another time measure, that of the total length of response, since it is possible that there is systematic variability between sentence types in, for example, pausing during the response itself. Particularly with more difficult three-term sentences such as those used in the present study, different results might be expected between these two time measures.

Previous studies have usually utilized a recall task requiring either repetition of the events of the stimulus sentence in "real time" order (e.g., Smith & McMahon, 1970; Clark & Clark, 1968) or, with young children, acting out the events of the sent-

ence with models (e.g., Amidon & Carey, 1972; Friedman & Seely, 1976). In addition to a recall task, we also used a card arrangement task, requiring the subject to manipulate cards, each with one clause printed on it, to indicate the real time sequence. It was hoped here that, by examining the first card used, we might be able to infer more directly how the sentence was comprehended--the "strategy of comprehension."

Adults, rather than children, were used as subjects in this study, in part due to the greater difficulty of the stimulus sentences. However, with Werner & Kaplan (1963), we assume a developmental continuity, in the sense that "objects of cognition" can be seen as resulting through a microgenetic analysis. In the case of language comprehension, this suggests a formal parallel between adult performance and prior developmental stages. For novel or difficult tasks, comprehension is slower, and therefore more easily observable.

The subjects in this study were 20 undergraduate women, serving as volunteers. Each subject was tested individually and performed both a verbal recall task and a card arrangement task; the order of tasks was counterbalanced. The verbal recall task required the verbal statement of the correct sequence of three-term sentences. For the card arrangement task, on each trial, three cards, each representing one of the three events, were presented in a vertical array with the first clause mentioned in the stimulus sentence as the topmost card. Arrangement of the cards was performed in a left-to-right, horizontal array accord-

ing to the real time sequence of events. For all trial errors were recorded. Two time measures were taken: response latency (that is, time from the offset of the stimulus sentence to the beginning of the response) and response length (that is, time from the beginning to the completion of the response). Additionally, order of selection of clauses was noted for the card arrangement task.

The stimulus sentences were the same for both tasks and consisted of eight possible combinations of three events in a sentence. There are four possible combinations of the temporal indices "before" and "after" within a sentence (i.e., before-before; before-after; after-before; after-after). Each of these four pairings was used with the main clause in the first position (e.g., "A before B and before C", etc.), and with the main clause in the last position (e.g., "Before B and before C, A", etc.), thus yielding the eight stimulus types. The "baseline" sentences, which paralleled the orderings of events in the stimulus sentences, used the words "first", "second", and "third" to indicate the temporal locus of clauses; the six possible orders of three events in real time were thus generated (e.g., "First A, second B, third C"; "Second B, first A, third C", etc.) Two exemplars of each of these eight sentence types and six baseline sentences were used in the final lists, thus yielding a total of 28 trials for each of the two tasks. Fourteen different sentences (8 sentence types plus 6 baselines) were used as a warm-up before each task. The stimuli were presented in a fixed random order,

and all sentences were "reversible"--that is, had no semantic indicator of temporal order (e.g., a sentence like "She got up, got dressed, and ate breakfast" was not used).

Separate analyses were performed on the stimulus sentences and the baseline sentences for each of the two tasks. Two two time scores were examined via analysis of variance, with Duncan's Multiple-Range procedure used to compare means. Errors and order of card selection were analyzed nonparametrically where that was possible. The response length measure did not yield significant differences between sentence types.

Results indicated that sentences with the same temporal indices (i.e., before-before; after-after) were easier to comprehend--that is, they produced fewer errors and shorter latencies--than sentences with different temporal indices (i.e., before-after; after-before). For sentences with the same temporal indices, the order of occurrence of the two events in the subordinate clauses is the same with respect to the event in the main clause. To produce a correct response, only the main clause need be transposed, while the other two events may be treated as a unit. It should be noted that in such sentences (e.g., "Before B and before C, A") the two events in the subordinate clauses are actually of indeterminate or ambiguous order with respect to each other, but not with respect to the event in the main clause. Results suggested that these clauses themselves were comprehended (i.e., repeated) in terms of their order of mention. For sentences with different temporal indices, each of the two

events in the subordinate clauses must be ordered differently with respect to the event in the main clause, requiring two transpositions.

How these more difficult sentences were comprehended was examined further in terms of the order of selection of cards in the card arrangement task. In this kind of task, it might be assumed that the clause selected first serves as an "anchor" for the ordering of the other two events with respect to it. It was found that the main clause was selected as the first to be ordered in 51% of the trials, suggesting that the main clause was used as an "anchor" with respect to the subordinate clauses. However, this finding resulted primarily from those sentences in which the main clause was the first-occurring event: for 98% of the "A before B before C" sentences and 92% of the "Before B, before C, A" sentences, the main clause was manipulated first in this task. For the other sentence types, in which one of the subordinate clauses represented the first-occurring event, no clear pattern of clause selection was apparent, thus suggesting that a main clause-subordinate clause "strategy" was not one of general applicability, at least for these types of sentences. It is possible, of course, that for some sentence types, subjects may have ordered the events covertly before selecting the first card, while for other sentence types part of the ordering task itself occurred during the card manipulation. However, analyses of the response length scores suggested no apparent pattern here.

Results suggested it was not generally the case that sent-

ences in which the main clause preceded the subordinate clauses were easier to comprehend than sentences with the main clause at the end. When the temporal indices were the same, sentences with the main clause last were as easy to comprehend as those with the main clause first. Similarly, an "order of mention" hypothesis did not fully account for the findings. The sentence which would be expected to be the easiest by application of both these principles (i.e., "A before B and before C") was found to be relatively easy on most measures. However, sentences in which the last event mentioned was the first to occur in real time (i.e., "Before B and before C, A"; "B before C and after A") were found to be equally as easy.

Similar findings were noted for the baseline sentences. Correspondence between order of mention and order of occurrence of events did not always result in easier comprehension. Here, the sentence expected to be the easiest (i.e., "First A, second B, third C") was found to require the shortest latency on the card arrangement task (although no different from the "Third C, second B, first A" sentence). But this same sentence produced longer latencies in the verbal recall task than three other sentence types (i.e., "Second B, third C, first A"; "Third C, first A, second B"; "Third C, second B, first A").

It is noteworthy that for both tasks, presenting the events in reverse order seemed to be as easy as presenting the three events in correct real time order. The most difficult baseline sentences, as with the sentences using "before" and "after" as

temporal indices, were those requiring a transposition in which the event described by the first or last clause mentioned occurred between the events of the remaining two clauses--thus requiring a different ordering of each of these two clauses with respect to the first or last one.

The most consistent findings across measures and across tasks were seen in terms of the kinds of transpositions required to produce a correct response. However, there was also considerable lack of consistency in the obtained order of sentence difficulty between different sentence types depending on the measure and the task. Such inconsistencies between tasks and between measures (e.g., in terms of order of mention and derivational simplicity principles) suggest that caution be exercised in interpreting results of studies using only one task and measure. Further, inferring a "strategy of comprehension" from some measures of sentence difficulty (e.g., errors) with two-term sentences would seem to be of limited generality in accounting for results with three-term sentences. Certainly there are differences between three-term sentences and two-term sentences such that the results of this study may not be directly compared with the results of other studies using two-term sentences. However, with two-term sentences, the kinds of errors possible to make are relatively limited. To infer a "strategy of comprehension" on the basis of analysis of errors with such sentences may be invalid due to the limited types of errors possible to make. Certainly, with three-term sentences, a greater number of kinds of errors

are possible; and the results of this study suggest less consistency than one "strategy of comprehension", or consistent combination of comprehension strategies, would lead one to expect.

FOOTNOTES

1. Paper presented at the 48th Annual Meeting of the Eastern Psychological Association, Boston, April, 1977. The study reported here was P. Arns' 1976 Honors Thesis at Mount Holyoke College.

REFERENCES

1. Amidon, A., & Carey, P. Why five-year-olds cannot understand before and after. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 417-423.
2. Barrie-Blackley, S. Six-year-old children's understanding of sentences adjoined with time adverbs. Journal of Psycholinguistic Research, 1973, 2, 153-165.
3. Bever, T.G. The comprehension and memory of sentences with temporal relations. In G.B. Flores D'Arcais & W.J.M. Levelt (Eds), Advances in psycholinguistics. New York: North Holland/America Elsevier, 1970.
4. Bloom, L. Language development. In Horowitz, F.D. (Ed), Review of child development research, Vol. 4 Chicago: University of Chicago Press, 1975.
5. Clark, E. How children describe events in time. In Flores D'Arcais & Levelt, ibid, 1970.
6. Clark, E. On the acquisition of the meaning of before and after. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 266-275.
7. Clark, E. How children describe time and order. In C.A. Ferguson & D.I. Slobin (Eds), Studies of child language development. New York: Holt, Rinehart, & Winston, 1973.
8. Clark, H.H., & Clark, E. Semantic distinctions and memory for complex sentences. Quarterly Journal for Experimental Psychology, 1968, 20, 129-138.
9. Friedman, W.J., & Seely, P.B. The child's acquisition of spatial and temporal word meanings. Child Development, 1976, 47, 1103-1108.

10. Hatch, E. The young child's comprehension of time connectives. Child Development, 1971, 42, 2111-2113.
11. Shilkret, R., Viets, & Wiener, M. Effects of redundancy and intonational stress in facilitating children's comprehension of temporally-ordered events. Paper presented at the Annual Meeting of the Linguistic Society of America, San Francisco, December, 1975.
12. Smith, K.H., & McMahon, L.E. Understanding order information in sentences: Some recent work at Bell Laboratories. In Flores D'Arcais & Levelt, ibid, 1970.
13. Werner, H., & Kaplan, B. Symbol formation. New York: Wiley, 1963.

The facilitating effects of redundancy as a pre-condition for
mastering more complex verbal forms

Morton Wiener and Judith Cormier and Robert Shilkret
Clark University Mount Holyoke College

A model of language comprehension has been explicated in previous work which attempts to account for "success" and "failure" of comprehension of auditory and graphic language in terms of a match-mismatch analysis of (1) the requirements of the language task and (2) the response patterns available to the comprehender, each assessed independently by some common measurement instrument (e.g., Wiener & Cromer, 1967; Cromer, 1968; Oaken, Wiener & Cromer, 1971; Steiner, Wiener & Cromer, 1971; Shilkret & Wiener, 1972; Oaken, 1973; and Wiener & Shilkret, 1974). Up to this point, at least, all of the efforts have focused primarily on the issues of match versus non-match in terms of a current assessment of the response characteristics of the responder relative to the requirements of the instance to be comprehended. While this model has been shown to be heuristic and valuable for dealing with responder characteristics and/or communication features which could be said to account for success and failure, it has not yet come to grips with explicating conditions which account for change in response possibilities. That is, what kinds of components or conditions might facilitate the possibilities of a change so that subjects will respond correctly to a more complex form of language than the one(s) to which they now respond. This report is an attempt to explicate a set of

principles for change and explore them in one experimental paradigm.

In the most general sense we regard two criteria as essential for the possibility of "change" from an "old" form (less complex) to a "new" one (more complex). The two forms (i.e., "old" and "new") must be ordered by some "ideal" scale -- that is, it must be possible to assess both forms independently with the forms related ordinally (i.e., any "new" form includes as much complexity as the "old" form, plus one additional component or relation). The transition from "old" to "new" forms, or what we call "change", will be accomplished if an intermediate form can be constructed in which there are co-occurring characteristics of the stimulus which is redundant with the earlier forms and also carries the form to be learned in the redundant construction. Once the "new" intermediate form has been mastered with inclusion of the "old" redundant form(s), adequate response to the more complex form without the old forms can become possible. In the most general sense, then, the following can be used as a schematic representation of the condition for "change."

1. "Old" form (e.g., riding a tricycle)
2. Redundancy, "Old" form together with "new" form¹
(e.g., use of "training wheels to ride a bicycle)
3. "New" form without "old" form (e.g., remove "training wheels)

It should be noted that we regard the two broad criteria outlined above as necessary for "change." Whether these criteria

are also sufficient for "change" is, from out point of view, an empirical question. That is, given that these two criteria are met, we would argue that change is possible; whether or not change occurs must be assessed in terms of particular individuals (or groups) in particular contexts with a particular combination of redundant features employed. However, as a general approach, we will hypothesize that these criteria are sufficient for "change" in any particular study.

It should also be noted that this kind of model of change hopefully avoids the problems of many developmental models which presupposes "conflict" as a prerequisite for change (e.g., Turiel, 1969). The concept of "mis-match" as conceived here does not imply the necessity that the comprehenders indicate "awareness" of the mismatch. In fact, we argue that it is reasonable to assume that the comprehenders always "comprehend the world in their own terms." This assertion does not deny that in some cases the person may well indicate "awareness" of a "mismatch" between the "comprehension" and the particular requirements of the context at that time. However, we would assume that this awareness occurs only subsequent to some feedback upon a prior "comprehension." In this special kind of instance of feedback the information in the feedback would be expected to include a form with some component(s) which is(are) redundant with the individual's original "comprehension" of the situation (the "old") as well as some "new" information. From this point, out analysis of change would apply even in these special instances. More generally, the mis-

match, as conceived here, is solely a statement by an observer, and derived from independent analysis of the individual and the task in terms of the same "ideal scale."

Since this model of change requires scaling of instances, it seems important that we discuss some issues about our use of one "ideal scale." Our scale is one which assumes that event A' is taken to be more "complex" than event A, if A' includes the discrimination of A, as well as at least one additional discrimination, combination or relation. For example, if an event is a correlation or coordination of two components (X and Y), such an instance is taken to be less complex by this scale than, for example, the conditionality of the same two components (X if Y); i.e., conditionality assumes correlation as well as an additional relation of the necessity of X as a condition of Y. In the same way, other logical relations could be scaled for complexity (see Rubano, Wiener & Shilkret in this report for a fuller explication of the development of one particular kind of scale).

It should be noted that these principles of ordering (i.e., incorporation of elements, grouping of elements, and relations) which we have used to specify our ideal scale is used solely as a heuristic device here in an attempt to order coherently and economically, and predict phenomena which it describes (see Kaplan, 1966; Toulmin, 1953). Thus, such a logical scale is not considered as "inherent in" the events themselves, the comprehenders' manner of constructing events, or even the inter-

action between the two. It may well be that under certain conditions, for certain comprehenders, and/or for certain events, the empirically observed order of difficulty as indicated by a performance measure may be found not to correspond with the posited ideal order of complexity which we expect. Were we to find inconsistent results, we would be required to re-explore our original scale and consider modification of the scale and re-consideration of the logic.

A second kind of issue in the use of the proposed ideal ordering derives from the assumption of ordinality ingredient in the way we have defined the scale. While we posit an increase in complexity with the inclusion of one further discrimination requirement, and each addition is assumed to increase complexity, additivity (interval scale) is not a necessary assumption in our derivation of an ordinal scale. That is, while the same differences in scale values may be obtained at different points in the scale, we do not hold that the degree of change need be the same at the two points -- they are only different in an ordinal sense (i.e., greater or lesser than). For pragmatic concerns, however, we may in any one instance, depending on the range of complexity involved, posit an interval scale to test the general proposition.

In contrast to many other approaches using complexity to explore language comprehension behavior (e.g., using Chomsky's transformational grammar in deriving an ordering of difficulty), the model proposed here requires that any complexity scaling include at the very least: (1) syntax (i.e., principles for com-

binning components and associated variations in sentence form); (2) the interrelationships among words in a sentence which is similar to Rosenberg's (1968) concept of associativity, but which includes also knowledge of the relationship of the content of the sentence to other knowledge (see Cormier, Wiener & Shilkret in this report); (3) number of discrete components and/or length; (4) degree of correspondence between the "real" order of events and the linguistic order of the represented event; (5) implied requirements of relational terms (e.g., temporal-spatial sequence vs. causality; coordination vs. conditionality; and (6) explicitness of the communication features available to indicate variations in all of the above.

To explore this model of facilitating a change in language behavior, we have selected some components in language instance which when combined with other components can then be ordered relatively easily by the scaling principles noted above. While we have selected temporal connectives to exemplify this approach we could, in principle, used the same paradigm to investigate changes in response to other variations, e.g., changes from coordination to subordination; or different forms of "conditionality" such as "when," "if," and "unless." This paradigm, then, will be used to assess a child's comprehension of the temporal markers "before" and "after," when given several kinds of intermediate redundant "training" versus conditions without such redundant training.

Fortunately, there is a considerable literature in which children's response to "before" and "after" has been explored.

While much of this literature is concerned with issues and controversies about other models of comprehension, the contradictory and inconsistent findings make it possible for us to (1) identify many of our requirements for controls and specify the grounds for including certain variables in this study and (2) explicate some principles which could account for the contradictory findings.

In general, it has been noted that performance differences have been correlated with the following variations in methodological procedures: (1) whether a child is required to give a verbal response or to perform a sequence of action; (2) whether the child performs the required actions in contrast to having some other actor (e.g., doll) perform action; and (3) whether the task requires the child to perform two different actions or perform the same action twice. While each of the variables can be considered important contributions to performance outcomes in any particular case, our model makes it possible to account for these differences by analyzing them in some common terms, that is, the number of simultaneous discriminations a child is required to make in any task condition. For example, let us compare the number of discriminations required when the child must give a verbal response, as opposed to performing some action. For the verbal response instance, the child is required to (1) understand the utterance, (2) understand requirements of the question to be answered (e.g., "What happened first?") and (3) related the question to the utterance; for the performance instance the child is only required to understand the utterance, and the relationship

of utterance to task, without the intervening question which also requires a semantic transfer of action to words. Similarly, a task which instructs the child to perform two different actions in sequence requires the child to discriminate at least, (1) the two actions, (2) the sequence and (3) which action goes with which apart of the sequence. The task which instructs the child to perform the same action in sequence only requires the discrimination of (1) an action and (2) the sequence. To the extent that the number of discriminations required varies from task to task, we would expect differences in performance outcomes even in response to the same syntactic pattern.

Given the number of variables introduced by different investigations -- some of which increase the complexity of the utterance and some of which increase the complexity of the task -- it is not surprising to us that the findings have been "inconsistent" and the investigators continue with additional studies to try to resolve these inconsistent findings.

Since our concern in this paper is more with principles which can account for the occurrence of change, that is, a child can now respond correctly to a more complex form than has been responded to previously, neither the difficulty of one form relative to the other (e.g., before vs. after), nor many of the other issues raised by the many other investigations of temporal connectives are focal for us.

Given our concerns, a series of sentences were constructed which would require the children to perform two actions in the

sequence specified by the temporal connectives. The basic sentence types consisted of two commands which were adjoined with either "then," "before," or "after." To order the several possible sentences which could be constructed with these sentences required us to consider several components which from our perspective contribute to the relative complexity of these sentences:

(1) the semantic component, that is, the specification that the relationship between the two actions (A, B) is temporal (specified by "then," "before," or "after"); (2) the syntactic pattern which imposes constraints on which temporal relationship is correct (A then B, or B then A); and (3) whether there is correspondence between the sequence in which the events are to be performed and the sequence in which they occur in the sentence.

Further, since our model has no grounds for asserting any inherent differences in complexity in the temporal terms "before" and "after," that is, there is no unequivocal principles which we can specify now would make it possible to describe one term as more complex than the other, (in that both terms only specify a temporal relationship between two events), we are required to analyze other components in an utterance to decide on "relative" difficulty. In this case, it is the sentence type rather than the temporal connective which is analyzed as more complex or less complex. The combinations of two commands with these temporal markers result in the following basic sentence types: (1) Do A, then do B, (2) Do A before you do B, (3) Before you do B, do A, (4) After you do A, do B, (5) Do B after you do A. Fea-

tures which vary between these sentence types are (1) the syntactic pattern which determines the position of the connective in the sentence, that is, whether the connective appears at the beginning of the sentence or between the two clauses, and (2) the sequence in which A and B appear in the sentence. To the extent that we consider redundancy essential for change we would argue that when all other features are held constant the location of the temporal connective in the middle of the sentence is less complex in that this pattern is consistent with the less complex form "do A and then do B" or even the earlier form "Do A and do B." The sentence patterns with the temporal connective in the middle position (1, 3, 5) are redundant with patterns which have already been responded to correctly and assumed to require no new discrimination at this time. This same argument holds for the sequence in which the two events, A and B, appear in the sentence. In the form, "do A and then do B" the events appear in the sentence in the same order in which they are to be performed. Thus, all other factors held constant, sentence types 1, 2, 4, which show correspondence between the order in which events are mentioned and the order in which they are to be performed, are considered less complex (i.e., it is congruent with the patterns which has already been responded to correctly and requires no new discrimination). It should be noted that the sequence in which A and B are to be performed is the same for all sentence types, although the order of appearance in the sentence does vary.

The first sentence type, "do A then do B," is considered the least complex in that the child is only required to discriminate the correct order and only one order of performance is possible with this form. Unlike the other connectives (before and after), it is not possible to vary either the position in the sentence (that is, "and then" must appear between A and B) nor the order in which the events are to be performed (that is, they must be mentioned in the order in which they are to be performed). In contrast, the introduction of "before" or "after" as temporal markers make it possible to vary both the syntactic position and the linguistic ordering of the task to be performed. Given these principles, we would order the "before" and "after" series of sentences in the following way:

- | | |
|--------------------------|---|
| (1) Do A then Do B | least complex in this series |
| (2) Do A before you DO B | new temporal marker with 2 logical possibilities |
| (3) After A then B | new temporal markers, 2 logical possibilities, new position |
| (4) Before B, A | new temporal marker, 2 logical possibilities, new position, and inverted order |
| (5) B after A | new temporal marker, 2 logical possibilities, inverted order but with "old" position which requires special discrimination in this case |

In the instance (5) above, if children had not learned the position of "and" and "then" as temporal markers, the sentences "B after A" would be expected to show no more errors than "Before B, A," since the number of discriminations are constant. However,

having learned the original form of "then," with correspondence of linguistic and event order, it is expected that a child will introduce error trials if position is responded to. In all of the forms except (5), a child must discriminate new forms -- either new terms, new position or linguistic order. In "B after A," the child must do some "relearning," rather than new learning, that is, ignore an old response.

We are hypothesizing that mastery of a more complex form will be facilitated to the extent that intermediate forms are introduced which incorporate components which are redundant with the previously mastered less complex form. At this time it is important to note that redundancy can be specified in terms of other components (e.g., intonational stress, Arns & Shilkret, 1977). In addition, we can introduce other even more explicit temporal cues which are redundant with the information in the temporal marker by the use of words such as "first," "second," "last," etc. Thus, the sentence "Before you do B, first do A," can serve as a transition between earlier forms and the more complex "Before you do B, do A." Further, we are predicting that if the child masters the transitional form, the redundant temporal information carried by "first" may well become unnecessary.

If we now include each basic sentence type, and "transitional" forms of each sentence type (that is, with "first" as an additional temporal marker) and then order them by our analysis of complexity, we obtain the following series:

Before

1. First do A, then do B
2. Do A, then do B
3. First do A before you do B
4. Do A before you do B
5. Before you do B, first
do A
6. Before you do B, do A

After

1. First do A, then do B
2. Do A, then do B
3. After you first do A, do B
4. After you do A, do B
5. Do B after you first do A
6. Do B after you do A

In sum, having scaled these sentences, we can assess empirically the tenability of this posited order and hypothesize that (1) each of the sentences with "first" will facilitate the learning of the more complex type without "first" and that (2) a subject receiving the series in increasing order of complexity will master more complex instances (have more correct trials) than children who receive the series in a non-ordered pattern. The basis for the second hypothesis is that some of the components of the less complex forms (e.g., position of marker, order of items) which are redundant with more complex forms may not be discriminated without ordering.

As we noted above, much of the earlier research with these temporal markers has included other variables which may include further discrimination requirements for our subjects. Therefore, we have attempted to control some of these variables in our tasks. Since differences have been found when subjects perform the action or manipulated other objects (usually dolls) to perform the actions, we included both conditions. While we used multiple verbs (e.g., put, move) they all required the same action on the

part of the child.

Another task variable which as far as we know has not been explicitly considered is one that we have investigated under the rubric of familiarity or unfamiliarity with the contents of sentences used in the various studies (see Cormier, Wiener & Shilkret, in this report²). We would argue that an important source of variability in performance can be overlooked if the content differences between sentences such as (1) "put the bed in the bedroom"; (2) "put the bed in the kitchen, and (3) put the prime minister in this office," are not considered. The difference in content between these sentences can be considered in two ways: first, in terms of the interrelationship among the words within the sentence (e.g., Rosenberg's concept of relative associativity) and secondly, in terms of the relationship between the content of the utterance and the knowledge the child is likely to already have about that content (familiar and unfamiliar). Given these considerations, we would expect different error rates in response to familiar vs. unfamiliar sentences whether it derives from associativity or from other knowledge of the world, even though the syntactic patterns are identical.

Materials

The stimulus sentences required the child to place objects in specified locations in a particular sequence. The objects that the children placed were familiar household furniture items, playground items, tools such as shovels and rakes, animals, a car and a truck. The locations included all the rooms in a doll

house plus the garage, back yard, and sidewalk. The Fisher Price Play Family dolls -- the mother, father, boy and girl -- were used for the conditions which required a different agent of the action.

Methods

Construction of Stimuli

Each sentence instructed the child to place familiar objects in certain locations in a specified sequence. For example, a typical sentence, based on the first sentence type (do A, then do B) would be "Put the couch in the living room, then move the car in the garage." The verbs, objects and locations were combined in the experimental sentences in the following way.

1) Six different verbs, all of which required the same action, were used to reduce response set. The position of these verbs -- put, place, set, move, bring and carry -- in the first or the second clause was systematically varied and no verb occurred more than once in any three successive sentences. Their order of occurrence was the same for all experimental conditions (i.e., typical vs. atypical; ordered vs. non-ordered).

2) Objects and locations were paired for the typical series such that familiar objects were to be placed in the locations where they are usually found (car in garage, sink in kitchen, bed in bedroom, etc.). A second series was constructed with only one variation -- that the same objects in the typical series were to be placed in locations where they were not usually found (car in living room, sink in bedroom, bed in back yard, etc.). All

objects and locations were identical for the typical and atypical series; the pairing of objects and locations was the only variation.

3) The subject-as-actor vs. doll-as-actor condition was met by presenting half the subjects with sentences which had been modified by the addition of the names of the dolls (mother, father, Johnny, Susie) used to perform the actions. Instead of the child carrying out the actions, then, the child was required to "act out" the sentence with the doll as actor. For example, instead of the utterance "put the couch in the living room before you move the car to the garage," the child would hear "Mother puts the couch in the living room before Johnny moves the car to the garage." This addition of another agent required the child to discriminate an additional pairing not present in the original sentence.

Combining the six levels of complexity and the two temporal markers results in ten sentence types. Types I and II are common to both the "before" and "after" series.

	Before		After	
Type I	(1)	First do A, then do B.	(1)	same as before list
Type II	(2)	Do A, then do B.	(2)	same as before list
Type III	(3)	First do A before you do B.	(4)	After you first do A, do B.
Type IV	(5)	Do A before you do B.	(6)	After you do A, do B.
Type V	(7)	Before you do B, first do A.	(8)	Do B after you first do A.
Type VI	(9)	Before you do B, do A.	(10)	Do B after you do A.

For each type of sentence presented the "before" and "after" instances were systematically varied. For Type III the sequence was (3) before, (4) after, (4) after, (3) before; for Type IV

the sequence was (6) after, (5) before, (5) before, (6) after; for Type V the sequence was (7) before, (8) after, (7) before, (8) after; for Type VI the sequence was (10) after, (9) before, (10) after, (9) before. This procedure required 2 exemplars of each sentence type. For the ordered condition the sentences were presented Type I through Type VI; the non-ordered series was presented in the fixed sequence Type II, V, III, I, VI, IV. Thus, while subjects in this condition would receive III before IV, and V before VI, they did not receive the entire series in an ordered form. These orderings makes it possible to not only examine the effects of "first" as a source of redundancy, but also permits us to look at the effects of the series itself as an additional source of redundancy. That is, we can compare the level of complexity (i.e., which point on the scale) mastered by the group who received the ordered series with the group who received the non-ordered series.

Subjects

A total of 149 pre-school and kindergarten children ranging in age from 4:0 to 6:0 years of age participated. The 92 (47 males, 45 females) kindergarten children were students in the same public school in a suburban town. Letters were sent through the school to parents of all children to explain the study and obtain written permission for their child's participation. All kindergarten children with parental permission were included except those who had reached their 6th birthday at the time of the study. The same letter was also mailed to parents of children who would

enter this same kindergarten school the following year. Fifty-seven pre-school children (29 males, 28 females) with parental permission who had reached their fourth birthday were included. Children were distributed among the experimental conditions by equating for the sex of the child at each of the 5 age levels used (4:0-4:4; 4:5-4:8; 4:9-5:3³, 5:4-5:7; 5:8 up to 6:0).

Procedures

After the children were acquainted with the experimenter they were told that they did not have to participate if they did not want to and that they could stop at any time. Two children declined and one child discontinued in the middle of the task. These three children were in the kindergarten age group.

After the array of objects and the doll house were shown to the children, the experimenter spent as much time as was necessary to determine that the children could correctly identify all the locations in and around the doll house where they would be required to place objects. The children were told that they would be given two objects and would listen to the tape recorder to hear where they were to put the objects. When they indicated that they understood, they were given two practice trials to familiarize themselves with the task, and to test their understanding. The form of the practice trial sentences was "Put A here and put B here." No children failed these practice trials. Two children's protocols were not included in the analyses when it was learned that one was not yet fluent in English (second language) and the other was receiving special language tutoring.

Results

The protocols were scored in two ways: (1) first, for a Sequence score where a trial was considered correct if the actions were performed in the correct sequence and (2) secondly, for a Sequence-Location score, where a trial was considered correct only if the actions were performed in the correct sequence and the objects were put in the correct location. Given two trials per sentence type with "before" and with "after" a subject would receive a score of 0, 1, or 2 for any sentence type for the "before" sentences or for the "after" sentences.

A series of Analyses of Variance were computed for (1) the Sequence scores as well as (2) the Sequence-Location scores. These analyses included the following factors⁴: (1) Age with 5 levels, (2) Sentence Type with 6 levels, (3) Redundant versus Non-Redundant sentences with 2 levels (the presence or absence of "first"), (4) Ordered vs. Non-Ordered series with 2 levels, (5) Sex of subject with 2 levels, (6) Typical vs. Atypical location series with 2 levels, and (7) "Before" vs. "After" sentences with 2 levels. In addition the following factors also were analyzed for trend components: (1) Age and (2) Sentence Types.

While Age as a main effect was significant ($F_{4,129} = 2.83$, $p < .05$), this finding was significant for the Sequence-Location series only, and interestingly enough this effect was not found as a linear component. Examination of the mean number of correct trials per Sentence Type at the different age levels showed the

following distribution of means: $M_{4:2} = 1.1$, $M_{4:6} = 1.4$, $M_{5:0} = 1.3$, $M_{5:6} = 1.3$ and $M_{5:10} = 1.4$.

A number of significant effects were also found for Age in interaction with one or more of the several factors analyzed for the Sequence-Location scores. The only significant effect for a first order interaction was age with Ordered versus Non-ordered series as a linear trend component ($F_{1,129} = 4.58$, $p < .05$). Examination of this effect showed a linear increase of scores over age for the non-ordered condition, but no increase over age was evident for the ordered series. That is, age as a variable showed an effect only when the experimental variables (that is, Redundancy and Order, etc.) were absent.

In the analyses computed, like Age, the Sex of the subject was a significant variable as a main effect only for the Sequence-Location score ($F_{1,129} = 4.69$, $p < .05$), with the females having a higher mean correct score. Interestingly, when the sex of the subject was a significant variable in interaction with the other experimental variables, the effects were noted only as third and fourth order interactions; all of the interactions which were significant also included the variables of Age, and Typical vs. Atypical location.

For the most part all of the remaining significant effects for the other experimental variables show equivalent effects when the data are analyzed in terms of the Sequence score or Sequence-Location score (see Table 1). The most efficient way for us to report these many findings is in tabular form with some comments added about each effect and the specific relationship

noted.

Insert Table 1 about here

The means for the significant effects noted in Table 1 are as follows:

1. The presence of "First" increases correct scores over all trials: For Sequence $M_{\text{First}} = 1.56$, $M_{\text{Non-First}} = 1.45$; for Sequence-Location $M_{\text{First}} = 1.35$; $M_{\text{Non-First}} = 1.23$.
2. The tasks for the connective "Before" shows more correct trials than for the connective "After" over all trials: For Sequence scores, $M_{\text{Before}} = 1.63$; $M_{\text{After}} = 1.38$; for Sequence-Location scores, $M_{\text{Before}} = 1.37$, $M_{\text{After}} = 1.21$.
3. Our subjects show significantly more correct trials for Typical location tasks: For Sequence scores $M_{\text{Typ}} = 1.58$, $M_{\text{Atyp}} = 1.42$; for Sequence-Location scores, $M_{\text{Typ}} = 1.45$, $M_{\text{Atyp}} = 1.13$. With rare exceptions, the correct scores for Typical locations were higher than for Atypical locations, over all of the factors, or with any or all other factors held constant.
- 4(a & b) Our subjects show fewer correct trials as the level of complexity of the sentence increases: Sequence scores, $M_{1+2} = 1.73$, $M_{3+4} = 1.55$, and $M_{5+6} = 1.24$; Sequence-Location $M_{1+2} = 1.53$, $M_{3+4} = 1.36$, and $M_{5+6} = .98$.
- 5(a & b) Our subjects show significant differences for Sentence Type and these effects show a significant linear trend: Sequence scores, $M_1 = 1.74$, $M_2 = 1.71$, $M_3 = 1.56$, $M_4 = 1.52$, $M_5 = 1.37$, $M_6 = 1.10$; Sequence-Location scores, $M_1 = 1.53$, $M_2 = 1.54$,

Table 1

List of Significant Effects ($p < .05$), Including Trend Components, in the Analyses of Variance of (1) mean correct Sequence scores, and (2) mean correct Sequence-Location for the experimental variables

Effect	Sequence F	df	Sequence & Location Combined F
1. Presence of "First" in Sentence	14.80	1/129	14.77
2. Before vs. After	71.96	1/129	30.41
3. Typical vs. Atypical Location Task	9.55	1/129	22.01
4a Comparison of Level of Sentence (Sentence Type 1 plus 2, vs. 3 plus 4, vs. 5 plus 6)	28.96	2/258	43.74
4b Linear Trend effects for this Analysis of Level	56.50	1/258	83.78
5a Comparison of the 6 types of sentences Main effect	7.54	2/258	5.22
5b Linear Trend	12.25	1/258	10.59
6. Effect of Ordering sentences are different for Redundant and Non-Redundant sentences	7.00	1/129	N.S.
7a Ordering of sentences show differences in correct scores for the sentence levels 1 plus 2, vs. 3 plus 4, vs. 4 plus 5	5.50	2/258	4.50
7b This effect is non-linear	10.22	1/258	8.92
8a Ordering seemed to facilitate the response to "After" sentences relatively more than for "Before" sentences	11.99	1/129	9.20

Table 1
(Continued)

Effect	Sequence F	df	Sequence & Location Combined F
8b This same effect was evident when sentences were of different levels of complexity	21.51	2/258	10.25
8c Linear effect	38.34	1/258	8.44
8d The differences between "Before" & "After" changed with the introduction of ordering and level of complexity of sentence	4.09	2/258	4.66
8e The effects were different as well for the presence of First vs. Non-First in combination with the level of complexity of the sentences	9.95	2/258	2.30 (N.S.)
8f Linear effect	8.25	1/258	2.70 (N.S.)
9. If one looks at all of the experimental variables, there appear to be a combined interactional effect such that with rare exceptions, the Atypical location had a lower mean correct score	6.05	2/258	3.27

$M_3 = 1.42$, $M_4 = 1.30$, $M_5 = 1.10$, $M_6 = .86$.

6. Just as Ordering has been shown to facilitate the correct response to the more complex sentences (see 5a & b above), Ordering seems to facilitate correct scores for the Non-redundant sentences over trials (that is, with first absent): For Sequence scoring:

- (1) $M_{\text{Ordered and Redundant}} = 1.56$
- (2) $M_{\text{Non-ordered and Redundant}} = 1.53$
- (3) $M_{\text{Ordered and Non-redundant}} = 1.55$
- (4) $M_{\text{Non-ordered and Non-redundant}} = 1.37$

While the correct scores for Sequence-Location show the same distribution, these differences were not significant.

7. The effect of Ordering shows a significant difference in the number of correct trials for the different levels, when Redundant and Non-redundant sentences are combined:

- (1) $M_{1+2 \text{ ordered series}} = 1.74$
- (2) $M_{3+4 \text{ ordered}} = 1.71$
- (3) $M_{5+6 \text{ ordered}} = 1.19$
- (4) $M_{1+2 \text{ non ordered}} = 1.72$
- (5) $M_{3+4 \text{ non-ordered}} = 1.38$
- (6) $M_{5+6 \text{ non-ordered}} = 1.28$

While the Non-ordered series scores appears "higher" for the Sentences 5+6, when one looks at the means of the "first" and the non-first sentences separately, we find that all of the effects can be attributed to presence of "first" in Sentence Type 5 in the non-ordered series.

8. Interestingly, while we have noted that our subjects showed higher correct scores for "Before" sentences than "After" sentences, this effect appears to differ depending upon the presence or absence of some form of redundancy, that is, first present or absent, Ordering vs. Non-ordering, or their combination. In general, an examination of Sentence Type shows a linear effect as predicted. However, if one explores the "Before" and "After" sentences in terms of Sequence scores at least, the one sentence "Before you do B, first do A" in the Non-ordered series has a greater number correct, then any other sentence in either Ordered or Non-Ordered conditions, including "First do A, then do B." This difference was less marked for the Sequence-Location score, where the sentence was the 5th highest score.

Discussion

In the Introduction above, we posit that it is possible to scale the amount of complexity in terms of the temporal connective, the position of this connective in the sentence, and the order of occurrence of the events in the linguistic form. Our data indicate that the mean number of correct trials per sentence type appear to follow the predicted scale. The effect of addition of the explicit cue of order (i.e., First), seems to be correlated with an increase of correct responses, particularly for the more complex sentence types. The means for the sentence types, with and without "first" are: (1) "Do A then do B," $M = 1.71$, $M_F = 1.74$; (2) "Do A before B," $M = 1.73$, $M_F = 1.70$; (3) "After do A, do B," $M = 1.33$, $M_F = 1.43$; (4) "Before do B, do A," $M = 1.24$, $M_F = 1.64$;

(5) "Do B, after do A," $M = .92$, $M_F = 1.09$. It is noteworthy that the scale appears to be a non-interval scale, at least in terms of the noted performance scores.

While the scale does show a linear effect as predicted, it is interesting to note -- and from our perspective still a problem -- that the requirement to respond to "after" seems relatively more difficult for our subjects than to respond to "before". This effect is clearly evident in terms of the overall conditions. In addition, the redundancy in the form of "first" also appears to be responded to more regularly for the marker "before" than when "first" appears with "after".

As expected, those subjects who were required to respond with atypical location placements showed fewer correct scores in their responses to sequence than those subjects who were given typical locations. While it would be reasonable to expect an increase in location errors when atypical location is a requirement, the finding of an increase in Sequence errors may seem unexpected. However, this finding is consistent with our analysis which lead us to posit that when a child is required to make a new discrimination, one not made previously, such discriminations constitute an increase of complexity and inferentially from our perspective an increase in the difficulty of the task. Associativity and/or prior knowledge of the world would seem to be significant contributions to comprehension ease or difficulty, even for the same syntactic pattern. (See Cormier, Wiener and Shilkret for other findings consistent with this assertion.)

The findings concerning the question of "training" via the

introduction of a transition, redundant sentence, versus a condition where the test sentence followed the transition sentence after several intervening but non-systematically ordered trials show that the effects are more complicated than our main effects suggest -- that is, a significant effect of Order. A closer examination of the means for the various conditions that contribute to the overall effect, seems to indicate that there was little, if any effect for the least complex trials and a little effect for the most complex trials. Whatever effect is evident seems to be in the "intermediate" sentences. We would expect then that redundancy will be a factor only for those instances where the level of complexity of the task to be mastered is not much greater (in scale units) than the complexity already evident as mastered by the child. That is, we would expect the immediately prior transition sentence to have more of a facilitating effect when the child is "close" to mastering the more complex form, but not when it is of less complexity or of considerably more complexity than the child has already mastered. We will add our caveat here then that this perspective always requires us to define the status of the child in the same terms as the task to be mastered -- match mismatch -- and only when the discrepancy is relatively small will we expect redundancy effects.

It is clear that a rigorous test of our general hypothesis about the importance of transitional redundant trials requires further and more systematic testing. It is equally important that the various forms of "redundancy" be tested separately. For example, if one were to test for typical knowledge as a redundant

feature in the learning of more complex forms, then one may test such discriminations using pre-school subjects and everyday temporally invariant sequences as follows: One could offer one group of subjects training with "before" and "after" with invariant sequences (e.g., "Put the socks on before you put on the shoes;" or "Put the car in the garage before you close the door") versus trials with variable possible sequences (e.g., "Put the car in the garage before you get the pencil;" "Put the red mitten on before you put the green mitten on"). One group of children could have several training trials with each of the invariant sequence with "before" and "after" in both positions and with order of mention varied systematically and then the children could be tested on variable sequence sentences. The second group of children could have the same number of "training" trials with variable sequence forms and then as with the other group, tested with the same variable sequence series of tasks as the first group of children. Similarly, we can establish a training series which includes trials with "first" as the redundant marker and contrast the effects with a group who had an equal number of training trials without "first" present as the redundant marker. In each case, we can compare the effects separately for sentences using "before" vs. "after" as temporal markers. However, prior to any exploration of the contributions of redundancy, as noted above, it is important to test the levels which the subjects-to-be-trained have already mastered.

Given this paradigm we can also begin to test the relative difficulty of temporal markers other than "before" and "after"

where positions in the sentence can be held constant. For example, we could test relative importance of the word effects by looking at sentences such as "When the father sits down, the girl turns on the TV," versus "After the father sits down, the girl turns on the TV;" or "The girl turns on the TV when the father sits down" versus "The girl turns on the TV after the father sits down." We would predict that the "after" sentences will be "easier" in that in our terms there is some conditionality in "when" and is a more complex form in our analyses than "after," which is taken as a discrete time marker.

If our hypotheses about redundancy is shown to be tenable, a more extensive study can be planned to assess whether children can be helped to learn any content where the sentences are less complex relative to texts written in more complex form.

FOOTNOTES

1. This procedure implies that what is to be taken as "new" for a given individual (or group) may be determined only in conjunction with some independent assessment of his prior learning (i.e., the set of "old" forms he has mastered). In one sense, this criterion is similar to the one of "inner logic" between "developmental stages" Loevinger (1969) has discussed regarding Piaget's developmental theory. However, by adopting a somewhat "weaker" criterion for assessment of "change" (i.e., any additional component or relation) we hope to avoid the possible logical and empirical difficulties which may ensue from assuming a particular order in universal, as if "in" the organism (see earlier discussion of this issue). Additionally, we hold that this kind of model has allowed for, and will continue to allow for, more specific predictions concerning individual differences.
2. This variation was introduced after examining the results of some pilot work and finding differential response to identical sentence types. We found fewer errors with pairings such as "bed-bedroom," "refrigerator-kitchen" than with pairings such as "chair-kitchen," "table-bedroom." It appears that perhaps an additional source of redundancy was introduced when the pairing in the sentence was a typical pairing of that object and that location, in that the instruction to the child was redundant with other information the child had about these objects and locations. For example, the bed typically goes in the bedroom and the stove typically goes in the kitchen, whereas "chairs" and "tables" can

typically be found in almost any location. With the more typical instance, then, the child, upon hearing "stove" could probably guess which location would be correct; with the less typical instance the child must discriminate all the components in the sentence before responding.

3. This group covered an 8 month age range. At the time of testing it was not possible to match the remaining groups without selecting populations from both pre-school and school children whose ages overlapped.
4. Examination of the data for "doll" as agent versus the child as agent of the action showed no systematic patterns of differences. The data for the two agent conditions were combined and the scores were then analyzed.

Knowledge of Referent and Comprehension of Language¹

Judith Cormier and Morton Wiener

Clark University

Robert Shilkret

Mount Holyoke College

Under a very diverse set of concern -- e.g., psycholinguistics, developmental psycholinguistics, sociolinguistics, social discourse and pragmatics -- there has been a remarkable resurgence of investigations about features of language or of characteristics of users that can be said to contribute to the success or failure to produce, or to comprehend, language. While these approaches seem to differ with respect to basic concerns, conceptual foci, population and specific methodological procedures of investigation, one commonality at least is that all look at "variants," "differences," or "failure" in production or in comprehension, as a means of specifying features of language or characteristics of users essential for normative language behavior by members of a communication group. One aspect of language and communication behavior, however, which has not always been of concern explicitly is what we, for want of a better term, call knowledge of "referent," where the term "referent" includes objects, events, relationships, ideas -- in short, the content of the language and communication. That is, although we agree that "competent" language use requires that speakers "know" linguistics and non-linguistic features of the communication system and/or other social features (e.g., who speaks to whom in what ways), one additional, or perhaps even a prior requirement would seem to

be that language users also know something of the world which is the subject matter or content of the communication -- an aspect most of us often seem to take for granted. We are not suggesting here that investigators have overlooked what we take to be an integral part of language and communication behavior. In fact, we believe that investigators may well recognize this aspect, although not always explicitly, when the controls for such variables as age and socioeconomic status are included, and contents appropriate for such groups are selected. When content has been of concern (e.g., for sociolinguists), this emphasis on content appears to be to account for "failures;" explaining "failure" or difference however has not led to an explicit articulation of the integral importance of shared knowledge of the contents of communication in language use and comprehension as a general principle. What we are saying is that while all of us would generally acknowledge that one might not comprehend new language forms -- new vocabulary, for example -- we are less likely to recognize explicitly the contribution of the referent knowledge component except for the "unusual" cases (e.g., cultural "deprivation") or where "esoteric" materials are presented (e.g., theoretical physics for a psychologist, or psychology for a theoretical physicist).

To demonstrate the importance of the contribution of "referent knowledge" to comprehension of a language passage in a more "typical" instance,² we will control linguistic features and social factors which are frequently taken to be the source of comprehension differences. To control for "knowledge" of language forms

or "features" we will use language instances which are invariant with respect to linguistic forms, patterns, and vocabulary by constructing a passage which could appropriately describe different events without varying the linguistic forms. This language instance can then be presented to subjects who are identical with respect to social factors such as "cognitive competence," age, general knowledge of the world, educational background, and social group membership. The subjects could be presumed, however, to differ in terms of what they know about the two referent events or contents "described" in the passage. The variable of concern to us, then, is differential "referent knowledge" about particular events and its contribution in comprehension as a general principle.

The procedure used in this study is similar to that used by Bartlett (1932) in his provocative investigation of memory and, by Paul (1959) in his extension of Bartlett's concerns, and more recently by Bransford and Johnson (1972) in their investigations of memory "activation." While we use "recall" as a measure of comprehension we would find it difficult to attribute all of the effect to "memory," once we note (1) the occurrence of paraphrase (e.g., Bartlett), (2) the differential recall of "familiar" vs. "non-familiar" material (e.g., Paul), and (3) the evidence that inferences are generated when individuals are asked to recall or identify previously presented material (e.g., Bransford & Franks, 1971). Although findings like these have often been noted as support of a view of memory as a "reconstructive" rather than a "re-duplicative" activity, from our perspective it is as reasonable to

argue that the material is "transformed" during "input," in which case comprehension is focal, rather than subsequent to "storage," which was "only" duplicative. That is, the basis for assuming a particular locus of "transformation" -- between "input" and "output" -- is open to question.

In any case, our primary concern here is not with "locating" the transformation activity at some point in a sequence, nor with demonstrating that "memory" is not a reconstructive activity. What we do argue is that it is as reasonable to hold that comprehension is a "constructive" activity, as it is to argue that memory is "reconstructive." An assumption of identical "input" -- as though the comprehender were a computer, or a "receiver" of language -- is contrary to our view of comprehension of communication as an ongoing activity which incorporates some construction or integration of the language "input," while reading (graphic form) or listening (auditory form). In this context, we would hold that language users read or hear and comprehend a language instance differently to the extent that there is differential knowledge of the linguistic forms or the content.³ In sum, while we will require reproduction of material, we hold that what is at issue is not only memory, but comprehension as well and predict that, in this instance, comprehension will differ to the extent that there is differential knowledge about the content. The primary hypothesis of the present study, then, is that when subjects are presented with identical language instances but describing different events, they comprehend and recall more of the

text to the extent that there is previous "referent knowledge" about the event. In addition, it was hypothesized that subjects produce more distorted and incorrect recall to the extent that there is less "referent knowledge" about the event.

Method

Subjects and Procedure

The subjects were 70 sixth grade children taken from the 3 middle tracks in a school in which children are assigned to one of 13 tracks based on school achievement tests. They were randomly assigned to three independent groups except that males and females from each of the three tracks were distributed as evenly as possible. The instructions to each group were similar in that all children were told that they would hear a tape-recorded passage which they would be asked to recall afterwards. The passage was:

"It seems like a complicated procedure at first, but it becomes easier after it has been done a few times. Before the cycle is begun, it is necessary to sort the materials into piles, according to kinds and colors. These things are then added to the water in the machine, along with certain chemicals. It is important to know the kinds of chemicals to add, and at what stage they should be added. When they are actually added will depend on the kinds of materials in the machine, the condition they are in, and how you want them to look when finished.

The inside of the machine has a tub to hold the contents, which has small holes in the bottom, allowing the water and chemicals to pass through and mix with the contents. As the materials are pushed back and forth in a circular motion, the liquid drains through the holes, is pumped up through the pipes, and comes in through the top again. As it comes in, it passes through a filter, or screen type device, which collects the small particles. Often, the only thing left to be done is the drying, but sometimes some of the items need to have a special finish put on, either before or after drying."

One group was told the passage they would hear would be about doing laundry (the greater referent knowledge group); a second group was told that it would be about papermaking (the lesser referent knowledge group); a third group was given no information about the referent. This third condition was similar to that used in the Bransford and Johnson study in that these subjects did not know what the referent was, in contrast to the "papermaking" condition where they knew what the referent was, although we presumed they did not know much about it.

Immediately after the students heard their passage, they were asked to indicate on a seven point scale how easy or hard it was to understand the text, with 1 being very hard and 7 being very easy. They were then asked to write down as much as they could remember of the passage they heard. When they were finished, they

were instructed to read over what they had written and encouraged to add anything they thought they had left out.

Results

Scoring and Reliability of Scoring

The stimulus passage was divided into 27 phrases or thought units, and a subject's score was the number of thought units judged to be present in the written protocol, matched against 27 thought units. Each thought unit judged to be present in the protocol was again scored by judges for one of three types of recall: (1) exact recall or a close paraphrase, (2) less explicit but correct recall, (3) distorted, incorrect recall, and (4) for intrusions, or the introduction of material not presented in the original passage. For example, one of the 27 thought units was 'the tub has small holes in the bottom.' Either that exact phrase, or a close paraphrase such as 'in the tub are little holes' were scored as a Type 1 recall, in that 'holes' and 'tub' were not juxtaposed as in the original passage, but their presence indicated that the information presented in the original thought unit has been included. A phrase such as 'the water went in the holes because the machine was broken' was considered a distortion of the original thought unit, in that the subject recalled something about 'holes' and 'water' but distorted the relationship between these elements. Phrases such as 'parts of the machine are dangerous' or 'some of the machines are blue' were judged intrusions in that no such information was presented in the original passage.

Two judges decided on both the number of the thought unit recalled and the accuracy of the recall. Three hundred and twenty seven instances were agreed upon, both for thought unit and recall category. A third judge looked at (1) the instances of non-agreement -- 125 for thought unit, that is, one judge scored a thought unit while the other did not; or one judge included one thought unit in another while the other judge scored two separate units; and (2) the 94 instances for differential use of recall category, to determine whether there were any systematic differences of scoring which could be reconciled, that is, for instances where one judge consistently uses different criteria. For example, for one thought unit there were thirteen instances where one judge scored "anything" being put in the machine, as Type I recall, while the second judge required specification of "sorted things" to score it the same way. The reconciliation decisions of this third judge made it possible to add 156 items to the data pool. There were only 9 instances which the third judge discarded as unreconcilable. The final data pool, with agreement for thought unit and recall category, amounted to 483 instances.

Analysis of Data

The data for (1) length of recall, (2) Type I plus Type II recall scores,⁴ (3) Type III error recall, and (4) Type IV Intrusions were analyzed in three factor analyses of variance -- that is, three Referent Conditions, three tracks in school, and male and female subjects.

In terms of the length of protocol, significant differences between Referent Conditions were noted ($F_{2,52} = 6.34, p = .05$, $M_{\text{laundry}} = 76$ words; $M_{\text{papermaking}} = 53$ words; $M_{\text{no referent}} = 40$ words). In terms of Type I plus Type II recall, significant differences between Referent Conditions were noted ($F_{2,52} = 5.94, p = .05$, $M_{\text{laundry}} = 7.8$; $M_{\text{papermaking}} = 4.9$; and $M_{\text{no referent}} = 4.5$) as well as for the three tracks ($F_{2,52} = 4.86, p = .05$, $M_{\text{high}} = 7.1$, $M_{\text{middle}} = 5.7$; and $M_{\text{low}} = 4.4$).

For Error scores (Type III) significant differences for Referent Conditions ($F_{2,52} = 6.97, p = .05$, $M_{\text{laundry}} = .5$, $M_{\text{papermaking}} = 1.7$, and $M_{\text{no referent}} = 1.0$) and for males and females were noted ($F_{1,52} = 4.53, p = .05$; $M_{\text{males}} = 1.4$, $M_{\text{females}} = .8$). For Intrusion scores, Type IV, differences among tracks ($F_{2,52} = 4.63, p = .05$, $M_{\text{high}} = .8$, $M_{\text{middle}} = 1.7$, $M_{\text{low}} = 1.6$) and for the three-way interaction of track, Referent Condition, and male-female were significant ($F_{4,52} = 2.62, p = .05$).

The subjects' judgments of passage difficulty (assessed prior to recall) were also analyzed. Significant differences for Referent Condition ($F_{2,52} = 4.20, p = .05$, $M_{\text{laundry}} = 4.78$, $M_{\text{papermaking}} = 4.15$, $M_{\text{no referent}} = 3.70$) and for the three-way interaction of Referent Condition by track by male-female ($F_{4,52} = 2.68, p = .05$) were noted. A significant correlation between correct recall (Type I plus Type II) and estimate of passage difficulty was found ($r = .25, p = .05$).

Discussion

The data were quite consistent with the hypothesis proposed and showed significantly better comprehension (or recall) for the familiar referent condition. There are two obvious concerns that require discussion. First, one might argue that these findings are no different than those reported by Barlett, and by Paul. That is, children only "reconstructed" what they already knew, and for children in the "laundry" condition this reconstruction would be more likely to be consistent with our criteria for comprehension than for children in the other groups. While we have no direct evidence to refute a memory reconstruction "interpretation," the data which show that subjects assess the laundry condition as the easiest for understanding is consistent with our view that "difficulty" in comprehension was apparent to the subjects before they were asked to recall the material. It may well be that other methods of assessing comprehension by using a reading technique -- e.g., cloze procedures of leaving every 6th or 7th word blank and asking the child to fill in the space; or reading errors or time to read -- may separate the difficulties of comprehension of "input" vs. the problems of recovery or "memory." Unfortunately, each of these methods we have considered seems to present some methodological difficulties which would present similar concerns about whether "comprehension" can be separated from other variables in any method.

The second issue is that it may seem "intuitively" obvious that a text describing an event presumed to be less familiar,

even though the forms are identical, would be more "difficult." Taking this "intuitive" approach, one might say that the present findings are nothing more than a demonstration of the obvious. In answer, the point can be made that the use of the terms "familiar" vs. "unfamiliar" are only categories, without any inherent explanatory value. Thus, while there may be differential difficulty in comprehension when the content can be said to be in either of these two categories, there is no specification of the parameters involved in the relative "difficulty."

The problem becomes more interesting, for us at least, when we note that the "phenomena" at issue is not different in kind from more everyday instances of comprehension and non-comprehension of the same material, even for one individual. Consider, for example, when one re-reads Freud, Piaget, or any other author who has offered us "remarkable reformulations" for a second or third reading. Although there are no new linguistic forms and no new dictionary definitions we often will recognize that we understand something quite differently on the subsequent readings. For the non-referent condition, one might say that the problem here is similar to other everyday instances when one arrives in the middle of an ongoing conversation, and while the words and sentences may be understood the conversation is not. In this pattern, the "difficulties" may be compounded, depending upon whether one knew what the context was, plus whether one "knew" other things about the referent event.

Given our findings, there is a further pragmatic concern

which arises when one looks at the literature in which different kinds of language users -- children of different ages, individuals from different social or cultural groups -- are assessed for their "knowledge" of language or content. It is apparent that all too often the tasks include variations in both the language forms (e.g., see Bormuth, 1970) and the content. It would seem essential, then, that tasks for the assessment of linguistics competence be controlled for referent knowledge, particularly for groups who can be assumed to have different referent knowledge; and that any task used to assess comprehension or achievement be controlled with respect to the forms. When the contents and the forms both vary, one cannot be certain that any differences found are attributable to the forms, the contents, or some combination of these two. This concern about simultaneous variation of forms and content is relevant not only in assessing what has already been learned, but also with respect to the teaching of new content or of new forms. Specifically, it would seem to be essential when teaching new or additional content to use linguistic forms which have already been mastered; and when teaching linguistic forms and patterns to use content which is already known, if we are to identify and account for instances of failure.

Even those instances for which "appropriate" controls of forms and contents are used can present an additional problem in assessment, at least to the extent that "memory" is used as a measure of comprehension. That is, when we do find instances of failure, or less than total recall, the basis for deciding whether

to consider it an instance of non-comprehension, an instance of "failure" of memory, or some combination of the two, is not clear. For example, since it was possible for subjects in our "laundry" condition to reconstruct what they already knew or comprehend what was presented, how do we decide on one of the following possibilities when accounting for less than total recall: (1) they failed to reconstruct everything they already knew, (2) they did not comprehend everything that was presented, or (3) they did not remember everything they comprehended. Similarly, when referent knowledge is not available, thus reducing the possibilities for reconstruction, how do we decide whether the subjects did not understand what they heard, or did not remember what they had understood?

While the findings reported here can be taken as evidence for a relationship between previous knowledge and comprehension of "new" material, we find ourselves faced with the more basic issue of how one might account for the learning of "new" material or the "integration" of "new" material with what is already known. At first glance, one may appeal to some "interaction" with the object of knowledge -- at least for subjects in the "laundry" condition in the present study. However, it is not clear how this type of explanation can account for difference in what becomes newly comprehended when one re-reads Freud, "studies" a text, or "learns" any new material which can be said to deal with ideas and not objects or actions. The issue for us, then, seems to reduce to the question of "how one is to explain the 'learning' of anything new."

FOOTNOTES

1. We wish to thank Dr. Shannon Devoe for her many helpful comments. We also wish to thank Irving Agard, Superintendent, Sooren Kachadoorian, Principal, teachers and pupils of the 6th grade Beal School, Shrewsbury School system, for their enthusiastic cooperation and participation.
2. We recognize that the relationship between "referent knowledge" and language is not undirectional, that is, language can as well be a precursor or requirement for additional knowledge. However, for this report, we are emphasizing the importance of "referent knowledge" as a prior or additional requirement for successful comprehension of a language instance.
3. In fact, we would argue further that we would also expect that language instances are "constructed" or comprehended differently to the extent that there are differences in the requirements of the task. For example, one will "comprehend" differently when the instructions are to "read for recall" vs. when the instructions are to "tell me what you think of this passage.
4. The data were also analyzed for Type I recall and Type II recall alone. All of the effects noted below may best be attributed to Type I recall (verbatim or close paraphrase) scores. We are reporting Type I & II as the more conservative estimate of comprehension.

REFERENCES

1. Bartlett, F.C. Remembering: A Study in Experimental and Social Psychology. New York: Cambridge University Press, 1932
2. Bormuth, J.R. On the Theory of Achievement Test Items. Chicago: University of Chicago Press, 1970.
3. Bransford, J.D., and Franks, J.J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.
4. Bransford, J.D., and Johnson, M. Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 717-726.
5. Cormier, J.H. Comprehension of a Descriptive Passage and Pre-experimental Familiarity with Referent Events. Unpublished M.A. Thesis, Clark University, 1976.
6. Paul, I.H. Studies in Remembering: The Reproduction of Connected and Extended Verbal Material. Psychological Issues, 1959, 1, #2.

Modifications in three groups of adults' speech to children and effects on children's language¹

Robert Shilkret

Lynnwood Andrews

Morton Wiener

INTRODUCTION

As part of this project, one study was conducted examining the complexity of younger children's language (i.e., 4- and 5-year-olds). This study was undertaken as an exploration of the ways adults with different kinds of experience with young children modify their speech to them, and, correspondingly, the ways the speech of the child may change as a function of hearing different kinds of adult language.

The ability to modify one's speech to meet the particular linguistic capabilities of one's audience has been widely investigated, for example, with reference to the Piagetian question of the child's gradual movement away from egocentrism. Such research has suggested that, at first, children do not modify their speech for a particular audience, but as they develop, the ability to take characteristics of the audience into account increases. From a different perspective, some psychologists interested in language development have recently begun to investigate some of the more specific characteristics of mothers to their young children. Generally, it has been found that adults seem to modify their speech to children in very specific and consistent ways. Simple syntactic forms (e.g., active declarative as opposed to passive sentences; Pfeuderer, 1969), consistent

pausing only after complete phrases (Broen, 1972), a high percentage of content words (Taburet-Keller, 1972), and a relatively small vocabulary (Broen, 1972) are examples of some characteristics of adults' speech to children.

Although an increasingly large body of data has been gathered on the specific changes in adult-child language, almost no studies have addressed the question of what may account for those stylistic variations. Moerk (1974) has pointed out that the level of language competence of the listener will affect the degree to which the speaker will change his or her speech. Thus, adults speaking to older children seem to use a style of language which more closely resembled adult-adult conversation in contrast with their speech to younger children. Snow (1972) investigated the effect of the presence of the child upon modifications in adults preparing instructions for children. A child-absent condition was found to elicit more complex patterns of language than a child-present condition. Finally, a controversy has arisen over whether the order of emergence of forms in child speech bears any relationship to the frequency of those forms in adult-child speech. Data concerning types of sentences used seemed, in one study at least, to suggest no such relationship (Pfuderer, 1969); on the other hand, studies of the use of subject-verb versus verb-object combinations found such a relationship (Bowerman, 1973). Thus, situational characteristics (e.g., whether or not the child is present; developmental status of the child listener) appear to be influential in adult speech modification.

This study was an attempt, first, to address the general

question of what factors may account for the specificity of adult speech modification to children. It would seem that amount of contact with children would be important in affecting adults' communication skills, in that such contact would allow an opportunity to know better the specific linguistic needs of the child. It was felt that subjects who had little experience with children would not modify their speech in exactly the same fashion as experienced subjects. Furthermore, it was hypothesized that a group of highly-experienced subjects, mothers, would not necessarily evidence the same modifications to their own children as they would to strange children; that is, a related question addressed by this study was whether a great deal of experience with a specific child may lead to speech which may be uniquely suited to that child's needs.

These hypotheses were examined in a story-telling situation where subjects told stories representing cartoon strips to both an adult and a child listener. Since the adult-adult speech situations and the adult-child situations were identical, direct comparisons could be made between the audience conditions (i.e., a direct assessment of modification of speech was possible). The cartoon strips were modified from a study by Bennett (1974) of the use of verbal connectives by children. The strips were composed to facilitate the use of specific kinds of connectives. Language was analyzed with particular emphases upon the use of logical relational terms and intonational features of speech. For the first of these, logical relational terms, both syntactic and semantic attributes of language were combined in an overall

measure of the explicitness with which subordinating conjunctions were used. The criteria for scoring subordinations were based upon a logical scoring system developed by Rubano (1976), and reported earlier. Connectives were rated as to the explicitness with which they defined a logical relationship between the main and dependent clauses. In addition, connectives which imply only coordination (e.g., "and") were not scored as subordinations, as well as connectives which, had they been used differently might imply subordination, but which seemed to be used in a coordinate manner. Thus, for example, "but" in "All his friends have two-wheelers, but he only has a tricycle" would be scored as a subordination since it seems to express contrariness; on the other hand, "Even though he's on his tip-toes stretching, he still can't reach it but his friend Sandra comes along..." would not be scored as a subordination for the "but" (the second sentence would receive a subordination score for the "even though"). For the second of these, the intonational features of speech, stresses, pauses, and rate of speech were examined. Stress, for example, has been shown to be important in young children's comprehension of sentences expressing temporally-ordered events (Shilkret, Viets, and Wiener, 1976), as described elsewhere in this report.

METHOD

The subjects for this study included 32 adults, divided into three groups, and 32 children. The adult subjects, all middle-class and English-speaking, were composed of three groups: (1) 12 mothers of children who also participated in the study, (2) 10

Experienced female undergraduates who had completed a four-month course which required at least 4 hours per week assistant teaching in a nursery school, and (3) 10 Inexperienced undergraduates who had had little or no experience with children under ten years of age (e.g., had not taken the course; had no younger siblings). Mothers were chosen to participate whose first born or only child was between 4-0 and 6-0 years, and was enrolled in the preschool program.

The 32 children who participated in this study were enrolled in a preschool program, where the study was conducted. The child listeners for the groups of Experienced and Inexperienced college students consisted of 20 middle-class, English-speaking children, half boys and half girls, and half between the ages of 4-0 and 4-11 and half between 5-0 and 5-11 years. Age and sex of child were counterbalanced across groups of adult subjects. The Mothers, treated somewhat different procedurally and analyzed separately, spoke both to their own children (all of whom were first-born, half male, half female; half each of the two age groups as before) and to strange children (children of the other mothers in this group in a yoked-pair design, and counterbalanced for age and sex.)

Materials

The materials consisted of four five-frame cartoon strips modified from Bennett (1974), with a fifth cartoon strip used as a practice trial. The picture-frame stories may be described as follows:

- (1) A boy is looking wistfully at a bike in the window of

a store; he is shown mowing a lawn; a man hands the boy money; the boy drops his money into a piggy bank; the boy is shown carrying his bank into the bike shop.

- (2) A boy is shown playing with a ball near a vase of flowers; the vase lies broken, with flowers and water scattered on the floor, along with his ball; the boy is discovered by his mother hiding under the bed; the boy is taken to the broken vase, mother pointing angrily; the boy is shown playing with the ball again near a lamp.
- (3) A girl is shown next to a bowl of apples, with a number of apple cores scattered around, with the girl holding her stomach and with a grimacing expression; girl is in bed with books strewn on floor, mother administering medicine and sunlight streaming through the window; still in bed, girl looks through window and sees children boarding school bus; girl smiles as bus leaves; girl, in nightgown, is shown next to a full bowl of apples, smiling.
- (4) A man, looking frightened, encounters a tiger; man is shown running, encounters a large snake in his path; he runs back in direction of tiger, again encountering the tiger; he runs back toward snake, again encountering the snake; the man is shown, stopped and with arms outstretched and looking sad, with snake on one side and tiger on the other.
- (5) (Used for practice) A boy is shown sitting on a box, looking and pointing toward toys on a high shelf with

a large box on the floor nearby; boy stands under the toys trying to reach them and missing by a large distance, with the two boxes on the floor nearby; boy is shown standing on one of the boxes under the shelf of toys, reaching for them and still not high enough; larger girl is shown pointing to second box with boy looking at her while standing on original box; boy is shown reaching toys while standing on two attached boxes.

These stories were used to elicit connective terms in description. Thus, for example, a purposive or means-end relationship can be described in stories #1 and #5. The boy in story #1 can be said to have "wanted" a bicycle and in order to obtain that, made money mowing lawns. Thus, mowing lawns was a sufficient condition to obtaining the bicycle and was logically subordinate to the goal. The story of the girl who ate too many apples (#3) portrayed a series of events which can be described as causal, as well as a temporal sequence of events; a similarly-construed sequence is presented in story #2. The frames of the cartoon pictures were 3 x 4 inches; the frames were arranged in left-to-right sequence and mounted on a horizontal strip of stiff board.

Procedure

The general procedure for all subjects, children and adults, was to tell the story for each of the cartoon strips under several conditions. The experimental sessions were scheduled while the children were in school and took place in a small room in the

school.

First, all 32 children were pretested by a female adult experimenter. During a practice trial with one of the cartoons (#5), the child was questioned if he/she hesitated a long time, became distracted, indicated confusion, skipped a picture, or rendered an extremely brief description of a picture. An attempt was made to begin with general questions (e.g., "Can you tell me some more about this picture?") before asking more specific ones. The experimenter avoided the use of connectives in the questions. When the child appeared to understand what was expected, the four experimental cartoons were presented one at a time. The same guidelines for questioning pertained to these trials as to the practice trial; if the child gave a very brief description, questions aimed at elaboration were postponed until the child had finished the story. The children were also prompted during the stories with such phrases as "Go ahead", "Good", or, most frequently, "Um-hmm." Each child was presented with the following instructions:

I came here today to listen to how children tell stories. Later, I will also be listening to how older people tell stories. Have you ever seen a tape recorder? (If child was familiar with recorder, instructions were continued; otherwise, child was given a chance to say something and then have it played back.) So, here are some cartoons. (Place cartoon series in front of child.) Each one tells its own story. I would like you to look at all the pictures, think about what's happening, and tell me the story. If you can't figure out the picture, ask me. Let's try this one (the practice story). Start with this picture (extreme left) and go that way (extreme right). (After Practice) Now I will give you these cartoons. If you would like to stop at any time, tell me, and I'll take you back to the playroom. Look at the pictures, think about what's happening, and then tell me the story.

Order of presentation of the experimental cartoons was counterbalanced (1-2-3-4; 2-3-4-1; 3-4-1-2; 4-1-2-3). These pretest sessions for the children, designed to elicit a "base-line" measure of the child's language use, were administered within one week before a second experimental session. All sessions were taped for later transcription.

The second set of experimental sessions involved the adults as well as the children. The adults, tested individually by the same female experimenter, were presented the following instructions:

This is a study about how children and adults speak with each other, and how adults speak with another adult. I have sets of cartoon strips which you will see in front of you. Each one tells its own story. Your job will be to look over the pictures and then tell me the story. You may ask me anything which is unclear about them. You may embellish in any way you feel is appropriate, but please do not delete or change the basic plot of any of the pictures. When you tell the stories to me, please speak to me as you would to any other adult. All the stories will be recorded on this tape recorder. After you finish, I will bring in (child's name) and you will tell the stories to him/her. He/she will then tell the stories back to you. The child has his/her own set of pictures, which are the same as the ones you have.

The adults told each of the four experimental stories to the adult experimenter in counterbalanced order. (Adult-adult condition.) Then one of the children was brought into the room and introduced to the adult. As the experimenter and child walked from the playroom to the experimental room, the child was told that a college student had come to tell him/her some stories and that the child was to tell the stories back to the adult when the

adult had finished. The adult told stories to the child in the same order as she had to the experimenter, and these trials constituted the Adult-child condition. For all sessions, the adult and child sat facing each other at a table with a screen between them to block visual contact; thus, communication was restricted to the verbal channel.

For the Mothers group of adults, the initial story-telling to the experimenter was identical with the Experienced and Inexperienced groups of adults. However, for the Mothers group, instead of telling the stories to a single child in the second part of the session for the adults, they told two of the stories to their own child and two to a strange child. The instructions which the mother received were the same as to the college students; however, the last paragraph indicated that the mother was to tell two of the stories to one of the children and then listen to the child tell them back. Then the same procedure was repeated with a second child. Instructions to children in the Mothers group were modified accordingly. Age and sex of the child and order of presentation of stories were counterbalanced, as was whether the mother told stories to a strange child first or to her own child first. This was accomplished by having two female experimenters run two mothers simultaneously in two separate rooms with the two children of these mothers. When two stories had been told in each of the rooms, the mothers simply changed rooms.

Finally, two weeks after this second session for the children, a third session was conducted with all the children, each of them telling the stories to the experimenter, as in the pretest.

Results and Discussion

All data were transcribed for scoring. The adult data and the child data were analyzed separately; for the groups of adults, the Mothers group was analyzed independently of the groups of Experienced and Inexperienced college students, given the difference in procedure for that group (i.e., two stories told to own child; two stories told to a strange child).

Analyses of Adult Data

The adult data were examined in terms of several response measures: total number of words per story, rate of speech, number of pauses, number of stresses, mean message unit length, number of subordinations, and complexity score. Reliabilities of each of the measures were computed for two judges' independent scoring of the data; for the complexity measure, three independent judges were employed. For the two groups of college students, three-way analyses of variance were performed for several of the response measures, with repeated measures on two factors (i.e., Audience: child vs. adult; Cartoons: 4 strips) and one independent groups factor (i.e., Experienced vs Inexperienced adults). Comparisons between means were accomplished by Duncan's Multiple-Range procedure.

Discrimination between stressed versus non-stressed words depended upon individual subjects' way of speaking rather than upon some absolute criterion applied across subjects (e.g., loudness). Each transcription was marked by two independent scorers for those words which "stood out" from the flow of speech; agree-

ment on this measure between the scorers was 94%. Stresses were defined as those words agreed upon by both judges. The scoring procedure for the total number of words spoken for each cartoon was a simple counting of the number of words for each story. Discrimination of words from non-words was based on a set of criteria devised by Schiefelbusch (1963). Ratios were then computed to yield number of stresses per word (number of stressed words/total number of words) for each story told by each subject.

Analysis of variance yielded a significant effect for Audience, $F(1, 18) = 49.01, p < .001$, with children being addressed with more words² than the adult experimenter. The interaction between Groups (Experienced vs. Inexperienced) and Audience (Child vs. Adult) was significant, $F(1, 18) = 10.23, p < .005$, with all four means composing this interaction differing significantly from each other ($p < .005$). Here, both Experienced and Inexperienced adults stressed more in speaking to a child than in speaking to an adult, but this difference was far more pronounced for the Experienced subjects than for the Inexperienced subjects. Experienced subjects stressed significantly more to children than Inexperienced subjects did; but these Experienced subjects stressed slightly but significantly less frequently than Inexperienced subjects when speaking to another adult.

The Cartoons main effect was significant, $F(3, 54) = 5.31, p < .005$, with the Apples cartoon producing fewer stresses than each of the other three (Boxes; Bike; Ball). The Cartoons x Groups interaction, $F(3, 54) = 2.46, p < .10$, suggested that this was due to the Inexperienced subjects producing fewer stresses

for this particular cartoon.

Pauses were defined as any interruption in the flow of sound between two words. Acceptable inter-observer agreement could not be achieved on very short pauses; thus only definite or longer pauses were scored for this analysis, with 95% agreement between two independent judges, and a pause being defined as that agreed upon by both judges. As with the stress measure, ratio scores were derived for each story told by each subject for the pause measure (pauses/total number of words). A three-way analysis of variance, as before, indicated a significant effect for Audience, $F(1, 18) = 7.92$, $p < .01$, with both college student groups pausing more in speaking with an adult than in speaking with a child. The Cartoons main effect was significant, $F(3, 54) = 4.23$, $p < .005$, with Bike and Ball receiving more pauses than Boxes and Apples. Neither the Groups main effect nor any of the interactions reached significance in this analysis of pauses.

A rate of speech measure was derived for one of the cartoons, Bike. Since number of pauses could affect the overall time for telling a story, this was controlled in deriving a rate estimate for each subject in the two Audience conditions (i.e., to child versus to adult). Of the two conditions for each subject, the entire story containing the fewer number of pauses was first timed; then the story in the second condition was timed, but only up to the same number of pauses plus one for the other story. If the "to child" story, for example, contained 7 pauses and the "to adult" story 10 pauses, the entire "to child" story would be timed but the "to adult" story would be timed only up until the beginning of the eighth pause. Reliability of this timing pro-

cedure was .99. A two-way analysis of variance, with Groups (Experienced vs. Inexperienced) as an independent groups variable and Audience (Child vs. Adult) as a repeated measures variable yielded no significant main effects or interactions.

A three-way analysis of variance, with variables defined as before, was performed on the total number of words for each story spoken by each subject. The Audience main effect was significant, with both groups of adults speaking more words to children than to adults, $F(1, 18) = 31.67, p < .001$. The Cartoons main effect was also significant, $F(3.54) = 9.68, p < .001$, with the Bike story being told in the fewest number of words, significantly fewer than the Apples and Ball stories, which themselves were no different from the Boxes story in terms of number of words. No interactions were significant in this analysis.

A set of criteria for defining message units was derived, and from this a mean Message Unit Length measure was derived as a substitute for possible utterance and sentence length measures. A "message" was defined as a group of words connected by syntactic and logical relations which expressed a single idea. Frequently, the message unit was synonymous with the syntactically complete sentence; however, syntactic "correctness" was not a criterion for a message unit. Boundaries between message units were discriminated by sentence boundaries as well as logical boundaries. Thus, run-on sentences were broken into their individual components--e.g., (He was quite afraid of what happened) (and he ran into his room) (and hid under his bed) (and his mother came in) (and showed him the broken vase and flowers) .

In addition, conjunctions not explicitly defining a relationship served as boundaries--e.g., (He had boxes instead of chairs) (It was a very avant-garde room) (So he was sitting on one of the boxes...) . Here, "so" expresses coordination as opposed to subordination (e.g., in order that). Agreement between two independent scores on message units was 99%. Disagreements were resolved by discussion. Mean message unit lengths were derived for each story told by each subject, and these scores were analyzed by a three-way analysis of variance, as before. Only the Cartoons x Audience interaction was significant, $F(1, 18) = 6.18, p < .05$.

A scoring procedure for Subordinations was established based on the method used by Bennett (1974). A Subordination was scored only if a relative pronoun occurred with its clause and if a conjunction was used in such a way as to identify explicitly the relation between the main and subordinate clauses. Simple coordinations expressed by "and" and "and then", for example, were not scored as Subordinations. Likewise, "but" was not scored when used to express coordination (e.g., "His mother--was talking to him about it and he still was looking very, very sad--uh--but then I don't think he learned anything from his mother"). However, when "but" indicated contrariness it was scored as a subordination (e.g., "...and he tries to reach the toys again--but he still can't quite reach them"). Agreement between two judges in scoring this measure was in excess of 95%, with differences resolved by discussion.

Ratio scores (number of subordinations/number of words) were derived for each story told by each subject, and these scores were examined by a three-way analysis of variance, with variables de-

fined as before. The Groups main effect achieved significance, $F(1, 18) = 3.77$, $p < .05$, with Inexperienced subjects using fewer subordinations than Experienced subjects. The Audience x Groups interaction was significant, $F(1, 18) = 3.33$, $p < .05$, and the pattern of means indicated that the Inexperienced college students used more "simplified" language (lower subordination ratios) than the Experienced college students in speaking both to children and to another adult. The Inexperienced subjects did not differ significantly in the rate of subordination to children as compared with their rate to adults; the highest rate of subordination was achieved by the Experienced subjects in their stories to children, and this rate was significantly higher than the rate of subordinations used by Inexperienced adults in their stories to another adult.

For the subordination measure, the Cartoons x Audience interaction was significant, $F(3, 54) = 3.61$, $p < .05$. Post-hoc examination of means revealed three significant differences: the Bike story when addressed to an adult contained fewer subordinations than both the Boxes story to adults and the Ball story to children; in addition, the Boxes story when addressed to children elicited fewer subordinations than the Ball story to children.

In summary of these analyses thus far, comparing the Experienced versus Inexperienced college students, it might be said that these two groups were not significantly different in terms of several more "traditional" measures of language performance (e.g., rate of speech; total number of words; message unit length.) However, the Experienced college students used more subordinations in speaking with children than did the Inexperienced college

students. While both groups used more stresses with children than with another adult, this was far more pronounced for the Experienced subjects than for the Inexperienced subjects. In a sense, it may well be that the Inexperienced subjects used stress in speaking to children when such stress was not really needed (i.e., relatively simple sentences, with few subordinations), whereas the experienced subjects used stress to "compensate" for their increased use of subordinations.

The language of the Experienced and Inexperienced college students was analyzed in terms of "complexity" of connective terms, that is, in terms of the scaling which is one of the foci of this research project. For this scoring, three judges scored the data independently, deriving a maximum complexity score for each of the stories told by each subject. Of a total of 792 pairwise comparisons of judges' scores, in 398 instances the scores were identical; in 205 instances the scores differed by one; in 93 instances by two; in 47 instances by three; and in 49 instances by four or more. The score assigned for each subject for each condition was the highest complexity score agreed upon by two judges. All instances where there was a disagreement between any two judges by more than two complexity units were examined individually and discussed by the group of judges; the relatively few differences that could not be resolved were compiled and used to suggest modifications in the scoring procedure for subsequent studies using this scale.

A three-way analysis of variance of these complexity scores, with variables defined as before, yielded a significant Audience

main effect, $F(1, 18) = 5.36$, $p < .03$, with a child being addressed more complex language than another adult. The Groups x Audience interaction was significant only at the $p = .11$ level, $F(1, 18) = 2.82$. Comparison of means indicated that Inexperienced college students did not differ in complexity in speaking to a child versus an adult, whereas Experienced college students' complexity scores were higher when speaking to a child ($\bar{m} = 7.13$) than when speaking to an adult ($\bar{m} = 6.03$; paired $t = -2.84$, $df = 9$, $p < .02$, two-tailed test.) Thus, the results of this analysis parallel those for the subordination scores.

The language of the Mothers group was analyzed separately, given the difference in procedure for this group (i.e., two stories to own child; two stories to strange child.) T-tests for related means comparing Mothers' performance in the Audience conditions (i.e., to child vs. to adult) yielded the following results: Mothers stressed more to children, $t(11) = 79.06$, $p < .001$; used fewer pauses to children, $t(11) = 92.09$, $p < .001$; spoke more total words to children, $t(11) = 4.62$, $p < .001$; and used more subordinations to children, $t(11) = 45.18$, $p < .001$, than they did in telling stories to another adult. In terms of complexity scoring, Mothers used more complex language with children than with another adult, but this difference was reliable only at the $p < .10$ level, $t(12) = 2.16$.

T-tests for related means were conducted for Mothers speaking to their own children versus Mothers speaking to strange children. Mothers used more stresses to strange children, $t(11) = 33.16$, $p < .001$, than they did to their own children. Similarly, Mothers subordinated more to strange children than to their own, $t(11) =$

19.67, $p < .001$. In this sense then, the results for the Mothers paralleled those for the Experienced college students: both use a relatively high rate of subordination and combine this with a higher rate of stress. For the Mothers group, this pattern is particularly evident in their speech to strange children, as compared with their speech to their own child.

Independent analyses compared the two groups of college students with the group of mothers. It was found that Experienced college students paused more to children than Mothers did, $t(14) = 62.37$, $p < .001$; they also used more subordinations to children, $t(14) = 2.50$, $p < .05$; and used more subordinations to adults, $t(14) = 9.80$, $p < .001$, than the Mothers did.

The results thus far support the hypotheses that experience with children may account for the specificity with which adults modify their speech to both children in general, and, in the case of Mothers, to one particular child. Furthermore, it appeared that the amount and kind of experience with children which the adult has had influenced the extent to which speech was modified, as indicated by the response measures analyzed in this study.

With regard to the stress measure, all three groups of adult subjects modified their speech to children in the same way, using more stress in speaking to a child than in speaking to an adult. However, this modification was much less for the Inexperienced adults than it was for the Experienced adults. Since it cannot be said, of course, that the Inexperienced college students had never had contact with children or have no knowledge of the specific communication needs of young children, it is not surprising that the

Inexperienced subjects modified their speech to children to some extent at least. The important point, however, is that the Inexperienced subjects did not seem to appreciate the extent to which their language might need to be modified when speaking to a child. With regard to the subordination measure, and paralleled by the complexity measure, Inexperienced adults did not modify their speech for children, whereas Experienced adults did. It should be noted that the subordination and complexity measures here can be viewed as the degree to which relations between events are made explicit in language. In this sense, the higher subordination and complexity scores of the Experienced adults and Mothers, as compared with the Inexperienced adults, can be taken as these groups' providing greater explicitness when speaking to children.

Comparisons between the Mothers and Experienced subjects appeared to suggest that nursery school assistant teaching is a kind of experience which might lead to speech modifications that differed from those employed by mothers. To the extent that teachers spend a large amount of time giving instructions and explaining reasons to children, it seems reasonable that their speech contained a higher degree of explicitness of logical relations and more pauses. Although, of course, it is not certain that mothers spend less time explaining and giving instructions to their children than teachers do, perhaps mothers have a broader range of different kinds of interactions with their children than do nursery school teachers. Such differences in the specific kinds of experiences with children may lead to

specific kinds of modifications, depending on the situation, the kind of communication, and so on.

It has sometimes been thought that mothers understand their children's speech, which to another adult may sound unintelligible. The results comparing Mothers speaking to their own children versus Mothers speaking to strange children implies that mothers do seem to use a unique communication style when addressing their own children. Perhaps this results from a different understanding of their child's speech than would be evidenced by a strange adult. It should also be emphasized that when addressing their own children, the mothers in this study demonstrated lesser evidence of explicitness and emphasis than when addressing strange children. A further examination of the specific kinds of subordinations and the specific words stressed may uncover an even greater recognition of particular children's needs than indicated by the response measures analyzed here.

Two very interesting serendipitous findings involved the adult-adult speech and a difference in apparent difficulty found when comparing between Cartoons. (1) Experience in a nursery school appeared to affect speech to adults as well as to children by the Experienced subjects; they were more explicit in their speech than both the Inexperienced and the Mothers groups. Given that the study was conducted in the nursery school, it may be that the setting included a set of general situational cues, to which the Experienced college students responded more than the other two groups. (2) A possibility for further research is suggested by differences obtained between the Cartoons. One of

the cartoon strips (Apples) was described by many of the adult subjects as the most difficult; some of the analyses also indicated that this cartoon was atypical. This cartoon was seen to be difficult because of the possible ambiguity of the story line. A further study could examine how different groups of adults respond when the material to be communicated is ambiguous; in a sense, such a situation corresponds with another kind of "real life" situation of communication between an adult and a child.

In summary, the role of experience with children was found to influence the language modification skills of adults speaking with children. The amount and kind of experience with children (teaching versus motherhood; long experience with one child) affected the extent to which the response measures were modified. Also, the possibility that teaching experience may affect speech to adults, and that the ambiguity of the subject matter may influence language modification was suggested.

Analyses of Child Data

It will be recalled that the children in this study told the stories on three occasions: in a pretest session within a week before the experimental session; immediately after hearing an adult from one of the three groups of adult subjects (Experienced; Inexperienced, Mothers) tell the stories; and, finally, in a posttest two weeks later. Analyses of children's language focused on complexity scoring. Each story told for each cartoon during each of the three testings was scored for complexity by three independent judges. It was necessary to omit one child from some of the analyses due to incomplete data. As before,

differences between judges were resolved by discussion. From the original scorings, there were a total of 1185 pairwise comparisons between judges' scores. In 677 instances the scores were identical; in 288 instances the scores differed by one unit; 131 instances differed by two units; 68 instances differed by three units; and 21 instances differed by four or more units. As before, the complexity score analyzed for each condition was the highest score agreed upon by two of the judges. There was a correlation of .92 between complexity score and length of the scored utterance.

Separate analyses were conducted on the complexity scores of the children who had heard Experienced or Inexperienced college students tell the stories to them (the third group of children had heard their own mother tell them two stories and a strange mother tell them the other two). A $4 \times 3 \times 2$ analysis of variance was performed on the scores of the children from these two groups of college students; variables in this analysis were Cartoons (4 stories) and Testings (3) as repeated measures, and Groups (hearing Experienced or Inexperienced adults) as an independent groups variable.

The Testings $\times$ Groups interaction was significant, but only at the $p < .10$ level, $F(2, 32) = 2.67$. A Duncan's Multiple-Range procedure indicated that these two groups of children were the same in the pretest ($\bar{m}$ -experienced = 2.33; $\bar{m}$ -inexperienced = 2.39). However, in the immediate posttest, the children who had heard Experienced adults increased in the complexity of their own stories ($\bar{m}$ = 3.28), while the children who had heard Inexperienced adults

decreased significantly and by almost the same amount as the other children had improved ($\bar{m} = 1.83$). The differences between each of these groups of children during immediate posttest and their own baseline scores during the pretest were significant, as was, of course, the difference between the two groups during the first posttest. The improvement evidenced by the children with Experienced adults was not manifest during the two-week posttest ($\bar{m} = 2.14$); conversely, the decrement in complexity scores of the children with Inexperienced adults was not manifest during the two-week posttest ($\bar{m} = 2.67$). While the two groups were not significantly different at the two-week posttest, each had changed significantly from its score during the immediate posttest (all comparisons at $p < .05$ or better). Thus, in summary, children who had heard Experienced college students tell them the stories obtained higher complexity scores immediately afterwards and children who heard Inexperienced college students decreased in complexity scores; neither of these effects was permanent in that both returned to a complexity level no different from their pretest scores on the two-week posttest. Thus, it would seem that the differences between the Experienced and Inexperienced adults' language was reflected in differences in the language of children who heard them, but these differences were not of long duration.

In this analysis, Cartoons was also significant as a main effect, $F(3, 48) = 15.80$, $p < .0001$. This effect was due to the Boxes story, which yielded a higher complexity score ($\bar{m} = 3.67$) than either of the other three (Bike = 2.17; Apples = 2.11; Ball = 1.82).

Complexity scores of children in the Mothers group were analyzed separately. A two-way analysis of variance was performed on these scores, with Cartoons and Testings as repeated measures, as before. The effect of Testings was significant, $F(2, 20) = 7.18, p < .01$. A Duncan's Multiple-Range procedure indicated that the pretest ($\bar{m} = 2.14$) was significantly lower than the immediate posttest ($\bar{m} = 3.45$) as well as the two-week posttest ($\bar{m} = 3.25$); the two posttests were not significantly different from each other. Thus, like the children who had heard Experienced adults, those children who heard Mothers also improved (i.e., higher complexity scores on immediate posttest), but, in contrast, the improvement of children who had heard Mothers was also evident two weeks later.

Cartoons was also significant as a main effect in this analysis, $F(3, 30) = 9.41, p < .001$; again, this effect was attributable primarily to the Boxes story, which produced higher complexity scores ($\bar{m} = 4.18$) than each of the other three (Bike = 2.52; Apples = 3.18; Ball = 1.91).

A question might be asked whether the children's telling of stories differed for those particular stories told by his/her own mother, as compared with those stories told by a strange mother. Again, the analyses of Mothers' language to their own child versus language to a strange child differed on some of the response measures analyzed, although these differences were not clearly interpretable. Given the counterbalancing procedure used in this study, this question could be addressed. New scores were derived for the children in the Mothers group, a separate set of scores for those stories told by his/her own

mother and a second set for those stories told by a strange mother. Cartoons were counterbalanced over these two conditions (own versus strange mother). A two-way analysis of variance, with Testings and Mothers (own vs. strange) both as repeated measures, did not yield a significant difference for the Mothers variable.

Finally, an independent set of analyses was performed examining all three groups of children. Again, this analysis may be of questionable validity since the procedure for the children in the Mothers group was somewhat different from the children in both the Experienced and Inexperienced adult groups. The three-way analysis of variance, with variables defined as before, yielded results parallel to the original analyses. Of interest here was the Groups x Testings interaction, $F(4, 48) = 2.52$, $p = .053$. The comparison of interest indicated that the children who had heard their mothers were no different during the immediate posttest than children who had heard Experienced adults; that is, both groups improved (i.e., increased complexity scores) by the same amount (in fact, as reported above, the numerical means are almost identical.)

The results of this study seem striking in that, first, differences were evidenced between three groups of adults in talking to young children. The children who had heard Experienced adults and Mothers increased in complexity of language, perhaps "imitating" the greater complexity of language they heard. The greater number of stresses used by these two groups of adults perhaps indicated the particularly important words, words which were used by the child in retelling the story. Furthermore, the

children who had heard Mothers maintained their greater complexity of usage in telling these stories two weeks later. In contrast, children who had heard Inexperienced adults actually declined in complexity of language use immediately after hearing them, but there was a rebound to their original level two weeks later. These results can be seen as beginning to specify some features in the linguistic environment of the child which may facilitate use of increasingly more complex language forms.

REFERENCES

1. Bennett, J. Usage of verbal connectives by children in relation to age and contextual variation. Unpublished Master's Thesis, Department of Psychology, Clark University, 1974.
2. Broen, P.A. The verbal environment of the language learning child. American Speech and Hearing Association Monographs, 1972, 17.
3. Bowerman, M. Early syntactic development: A cross-linguistic study with special reference to Finnish. Cambridge, England: University Press, 1973.
4. Moerk, E. Changes in verbal child-mother interactions with increasing language skills of the child. Journal of Psycholinguistic Research, 1974, 3, 101-11.
5. Pfuderer, C. Some suggestions for a syntactic characterization of baby-talk style. The structure of linguistic input to children, Working Paper 14. Unpublished manuscript, Language-Behavior Research Laboratory, University of California, Berkeley, 1969.
6. Rubano, M. An analysis of the use of logical connectives by children. Unpublished manuscript, Clark University, Department of Psychology, Worcester, Mass., 1976.
7. Schiefelbusch, R. Language studies of MR children. Journal of Speech and Hearing Disorders Monograph Supplement, 1963, 10.

8. Shilkret, R., Viets, M., & Wiener, M. Effects of redundancy and intonational stress in facilitating children's comprehension of temporally-ordered events. Paper presented at Annual Meeting of the Linguistic Society of America, San Francisco, December, 1975.
9. Snow, C.E. Mothers speech to children learning language. Child Development, 1972, 43, 549-565.
10. Tabouret-Keller, A. L'acquisition du langage entre 2 et 3 ans. Bulletin de Psychologie, 1972-3, 26, 140-147.

FOOTNOTES

1. Grateful acknowledgement is expressed to Frances Kerr, Director, and to the staff of the Gorse Child Study Center of Mount Holyoke College for their cooperation in all aspects of this study. The adult language part of this study constituted Lynwood Andrews' 1976 Honors Thesis from Mount Holyoke College. Lynda Godkin assisted in running part of this study and in the scoring and analysis of data. Maureen Callanan, Anne Claunch, Lynda Godkin and Suzannah Heard, assisted in scoring and analysis of data.
2. Means are not reported for these analyses, since ratio scores were analyzed, and these are relatively meaningless for purposes of discussion. However, only differences which are taken to be significant are reported, with the level of significance indicated.

Additional Bibliography

- Amidon, A., and Carey, P. Why five-year olds cannot understand before and after. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 417-423.
- Bennett, J. A study of the usage of verbal connectives for children at different ages. Masters Thesis, in preparation, Clark University, 1974.
- Bernstein, B. Class, Codes, and Control. Vol. I. Theoretical Studies towards a Sociology of language. London: Routledge and Kegan Paul, 1971.
- Bregman, A.A., and Strasberg. Memory for the syntactic form of sentences. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 396-403.
- Cazden, C. The neglected situation. In F. Williams (Ed.), Language and Poverty. Chicago: Markham Press, 1970.
- Clark, E. On the acquisition of the meaning of "before" and "after." Journal of Verbal Learning and Verbal Behavior, 1971, 10, 266-275.
- Cole, M., and Bruner, J.S. Cultural differences and inferences about psychological processes. American Psychologist, 1971, 26, 867-876.
- Cromer, W. The effect of "pre-organizing" reading material on two types of poor readers. Unpublished Doctoral Dissertation, Clark University, 1968.
- Epstein, W. The influence of syntactical structure on learning. American Journal of Psychology, 1961, 74, 80-85.
- Epstein, W. A further study of the influences of syntactical structure on learning. American Journal of Psychology, 1962, 75, 121-126.
- Epstein, W. Some conditions on the influence of syntactical structures on learning: Grammatical transformation, learning instructions, and "chunking." Journal of Verbal Learning and Verbal Behavior, 1967, 6, 415-419.
- Epstein, W. Recall of word lists following learning of sentences and anomalous and random strings. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 20-25.
- Falmagne, R.J., ed. Psychological Studies of Logic and Its Development. Lawrence Erlbaum, Inc., in preparation.
- Fillenbaum, S. Memory for gist: some relevant variables. Language and Speech, 1966, 9, 217-227.

- Fillenbaum, S. On the use of memorial techniques to assess syntactic structures. Psychological Bulletin, 1970, 73, 231-237.
- Fodor, J., and Garrett, M. Some reflections on competence and performance. In Lyons and Wales (Eds.), Psycholinguistic Papers, Edinburgh: University Press, 1966.
- Foss, D.J. and Cairns, H.S. Some effects of memory limitation upon sentence comprehension and recall. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 541-547.
- Garrett, M., and Fodor, J. Psychological theories and linguistic constructs. In Dixon and Horton (Eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.
- Gleitman, L., and Shipley, E. The acquisition of linguistic structure: The emergence of the child as grammarian. Technical Report, Appendix D, Grant MH-20041, July, 1971.
- Glucksberg, S., and Danks, J.H. Grammatical structure and recall: A function of the space in immediate memory or recall delay? Perception and Psychophysics, 1969, 6, 113-117.
- Gottlieb, R., Wiener, M., and Mehrabian, A. Immediacy, relief-discomfort, quotient, and content in verbalizations about positive and negative experiences. Journal of Personality and Social Psychology, 1967, 7, 266-274.
- Greene, Judith. Psycholinguistics. Baltimore: Penguin, 1972.
- Gruber, J.S. Topicalization in child language. Foundations of Language, 1967, 3, 37-65.
- Ingram, D. Transitivity in child language. Language, 1971, 47, 888-910.
- Jancovic, M. Age-related changes in hand and arm movements as nonverbal communication. Unpublished Masters Thesis, Clark University, 1973.
- Kagan, J. Formal discussion. In Bellugi and Brown (Eds.), The Acquisition of Language. Society for Research in Child Development Monographs, 1964, 29, No. 1, 169-172.
- Kaplan, B. The study of language in psychiatry, in Arieti, S. American Handbook of Psychiatry, Vol. 111, New York: Basic Books, 1966.
- Kaplan, Elenor, and Kaplan, G. The prelinguistic child. In Eliot, J. (Ed.), Human Development and Cognitive Processes. New York: Holt, Rinehart, and Winston, 1971.

- Kelly, George A. The Psychology of Personal Constructs. New York: Norton, 1955.
- Kempler, B., and Wiener, M. Personality-perception: characteristic responses to available part-cues. Journal of Personality, 1964, 32, 57-74.
- Labov, W. The logic of non-standard English. In F. Williams (Ed.), Language and Poverty. Chicago: Markham Press, 1970.
- Lane, H., and Schneider, B. Some discriminative properties of syntactic structures. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 457-461.
- Loevinger, J. Theories of ego development. In Berger, L. (Ed.), Clinical-Cognitive Psychology, Englewood Cliffs, New Jersey: Prentice-Hall, 1969.
- Makita, K. The rarity of reading disability in Japanese children. American Journal of Orthopsychiatry, 1968, 38, 599-614.
- Markham, E. The facilitation of part-whole comparisons by use of the collective norm "family." Child Development, 1973, 44, 837-840.
- Marks, L.A., and Miller, G.A. The role of semantic and syntactic constraints in the memorization of English sentences. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 1-5.
- Martin, E., and Roberts, K.H. Grammatical factors in sentence retention. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 211-218.
- Martin, E., and Roberts, K.H. Sentence length and sentence retention in the freelearning situation. Psychonomic Science, 1967, 8, 535-536.
- Martin, E., Roberts, K.H. and Collins, A.M. Short-term memory for sentences, Journal of Verbal Learning and Verbal Behavior, 1968, 7, 560-566.
- Matthews, W.A. Transformational complexity and short term recall. Language and Speech, 1968, 11, 120-128.
- Mehler, J. Some effects of grammatical transformation on the recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 346-351.
- Mehrabian, A., and Wiener, M. Non-immediacy between communicator and object of communication in verbal messages. Journal of Counseling Psychology, 1966, 30, 420-425.
- Miller, G.A. Some psychological studies of grammar. American Psychologist, 1962, 17, 748-762.

- Miller, W., and Ervin, Susan. The development of grammar in child language. In Bellugi and Brown (Eds.), The Acquisition of Language. Society for Research in Child Development Monographs, 1964, 29, No. 1, 9-34.
- Morse, P.A. Speech perception in six-week-old infants. Paper presented at Society for Research in Child Development Meetings, Minneapolis, April, 1971.
- Oakan, R. Language comprehension and comprehension difficulties. Unpublished Doctoral Dissertation, Clark University. 1973.
- Oakan, R., Wiener, M., and Cromer, W. Identification, organization, and reading comprehension for good and poor readers. Journal of Educational Psychology, 1971, 62, 71-78.
- O'Connell, D., Turner, E., and Onuska, L. Intonation, grammatical structure, and contextual association in immediate recall. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 110-116.
- Orenstein, H.B., and Schumsky, D.A. The effects of sentence length and grammatical structure in a serial learning task. Psychonomic Science, 1970, 19, 243-244.
- Pepper, Stephen C. World Hypotheses. Berkeley: University of California Press, 1942.
- Perfetti, C.A. Sentence retention and the depth hypothesis. Journal of Verbal Learning and Verbal Behavior, 1969a, 8, 101-104.
- Perfetti, C.A. Lexical density and phrase structure depth as variables in sentence retention. Journal of Verbal Learning and Verbal Behavior, 1969b, 8, 719-724.
- Robbins, O. Regulator behavior in restricted and elaborated code using groups. Masters Thesis in preparation, Clark University, 1974.
- Rosenberg, S. Association and phrase structure in sentence recall. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 1077-1081.
- Rosenberg, S. Language habits and the recall of connected discourse in Eric M. Zale, Ed., Proceedings of the Conference on Language and Language Behavior. New York: Appleton-Century-Crofts, 1968.
- Rosenberg, S. Recall of sentences as a function of syntactic and associative habit. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 392-396.
- Sachs, J.S. Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 1967, 2, 437-442.

- Sahakian, William S. Use and abuse of logic in psychological literature. American Psychologist, 1973, 28, 942-944.
- Savin, H., and Perchonock, E. Grammatical structure and the immediate recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 348-353.
- Shilkret, R. The facilitative effects of some "melodic" features of speech in sentence processing. Unpublished Masters Thesis, Clark University, 1970.
- Shilkret, R., and Wiener, M. The contribution of syntactic and parasyntactic cues in the comprehension of spoken and written language. U.S. Department of Health, Education, and Welfare, Office of Education Grant No. OEG-1-70-0009 (509), February, 1972.
- Toulmin, S. The Philosophy of Science. London: Hutchinson University Library, 1953.
- Turiel, E. Developmental processes in the child's moral thinking. In Mussen, P., Langer, J., and Lovington, M. (Eds.), Trends and Issues In Developmental Psychology, New York: Holt, Rinehart, and Winston, 1969.
- Wearing, A.J. Recognition memory for sentences of varying length. Psychonomic Science, 1969, 15, 221-222.
- Wearing, A. J., and Crowder, R.G. Dividing attention to study sentence acquisition. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 254-261.
- Weeks, T.E. Speech registers in young children. Child Development, 1971, 42, 1119-1131.
- Weir, Ruth H. Language in the Crib. The Hague: Mouton, 1962.
- Werner, H., and Kaplan, B. Symbol Formation. New York: Wiley, 1963.
- Wiener, M., and Cromer, W. Reading and reading difficulty: A conceptual analysis. Harvard Educational Review, 1967, 37, 620-643.
- Wiener, M., and Devoe, Shannon. Parent's language usage patterns and child's school performance. Grant Proposal NICHD, 1974.
- Wiener, M., and Kleepies, P. Some comments and data on partial cue controversy and other matters relevant to investigations of subliminal phenomena: a rejoinder. Perceptual and Motor Skills, 1968, 27, 847-861.

Wiener, M., and Mehrabian, A. Language Within Languages: A Channel In Verbal Communication. New York: Appleton-Century-Crofts, 1968.

Wiener, M., and Schiller, P. Sublimal perception or perception of partial cues. Journal of Abnormal and Social Psychology, 1960, 61, 124-137.

Wright, P. Two studies of the depth hypothesis. British Journal of Psychology, 1969, 60, 63-39.

Yngve, V.H. A model and an hypothesis for language structure. Proceedings of American Philosophical Society, 1960, 104, 444-466.